

DOCUMENT RESUME

ED 041 643

PS 003 671

TITLE	A Comparative Study of the Impact of Two Contrasting Educational Approaches in Head Start, 1968-69.
INSTITUTION	Bank Street Coll. of Education, New York, N.Y.
SPONS AGENCY	Office of Economic Opportunity, Washington, D.C.
PUB DATE	69
NOTE	68p.
EDRS PRICE	FDRS Price MF-\$0.50 HC-\$3.50
DESCRIPTORS	*Achievement Gains, Child Development, Cognitive Ability, *Comparative Analysis, Curiosity, Learning Readiness, Measurement Instruments, *Preschool Programs, Program Descriptions, *Research Problems, Self Concept, *Test Results, Verbal Communication
IDENTIFIERS	Project Head Start

ABSTRACT

As part of a national evaluation of Head Start, a comparison of school readiness and childhood development approaches to preschool education was attempted, but major methodological problems were encountered. It was not possible to find the study samples called for in the original plan, i.e. a child-readiness program of the Bereiter-Engelmann type, and a child development program that was a suitable example. A compromise selection of two Head Start centers included one that was child development-oriented, and one that had a modified Montessori program. A comparison sample was selected from a middle class child development-oriented private nursery school. The children were pre- and posttested on measures of cognitive skills, curiosity, self-concept, and spontaneous language. Individual child observations were also made. However, the original data collection plan was severely curtailed because of lack of time and testing space. The results of the study are not definitive but indicate that the middle class children were more able to benefit in demonstrable ways from a year of preschool education. However, the private program was judged to be of much better quality than the Head Start programs in the study. (NH)

F OEO

U. S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

PS

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE
PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS
STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION
POSITION OR POLICY.

Head Start Evaluation and Research Center

A COMPARATIVE STUDY OF THE IMPACT OF TWO CONTRASTING
EDUCATIONAL APPROACHES IN HEAD START

1968-69

Research Division
Bank Street College of Education
216 West 14th Street
New York, N.Y. 10011

ED041643

PS 003671

Staff

Herbert Zimiles, Ph.D., Director

Doris Wallace, Coordinator

Marcia Judson, B.A., Research Assistant & Center Liaison

Harvey Asch, Ph.D., Research Associate

Lois Chaffee, B.A., Research Assistant

Robert David, M.A., Research Assistant

Martha Friedrichs, B.S., Research Assistant

Anne Gordon, B.A., Research Assistant

Lyn Gordon, Research Assistant

Dinah Heller, M.A., Research Assistant

Judith Karelitz, B.A., Research Assistant

Kathryn Kirshner, B.A., Research Assistant

Katherine Oram, Field Staff Assistant

Janna Roth, B.A., Research Assistant

Jacqueline Royce, Ed.M., Research Assistant

Shirley Smith, Research Assistant

Rebecca Wallace, Field Staff Assistant

Carol Carr, Secretary

Maggie Newman, Secretary

Table of Contents

	<u>Page</u>
Introduction	1
Description of Centers	5
The Instruments of the Study	9
Results	11
Discussion	30
Appendix	

A Comparative Study of the Impact of
Two Contrasting Educational Approaches in Head Start

This study is a direct outgrowth of several years' participation in the national evaluation of Head Start; it served as a vehicle for attacking some of the major methodological problems that were identified in previous evaluation studies. As originally conceived, the study was designed to accomplish the following: (1) to describe in greater detail the differences between two contrasting approaches to early childhood education -- the School Readiness approach and the Child Development approach -- by studying the learning environments created by these two orientations; (2) to identify those dimensions of child behavior which are relevant to the study of pre-school influence by providing a record of behavioral change through intensive naturalistic observation of individual children during the course of a year of preschool education; (3) to assess the effects of two contrasting modes of preschool education on the participating children along a set of selected dimensions by means of pre-post test comparisons; and (4) to compare the functioning of Head Start children with a group of middle-class children of the same age who were also attending preschool.

The design of the study arose out of a number of methodological considerations: (1) Previous efforts of evaluation studies to differentiate teaching methods observed in various classrooms failed to arrive at a reliable method for distinguishing among naturally occurring variations in teaching behavior; nor did these observable variations in teaching behavior bear a consistent relation to particular outcome variables. It seemed prudent, therefore, to identify significant variations in teaching strategy explicitly formulated by educators and deliberately adopted by specific teachers and to make a comparison of these methods as enacted by their exponents, rather than to trust to the modest and diverse levels of variation which are obtainable from random selection of study classrooms; (2) because of the greater precision and economy of measurement afforded by psychological testing methods, most assessments of the effect of Head Start on participating children have taken this route despite the lack of both validity and relevance of most such measures. It was time to overcome the rigidity of method which seemed to have permeated all assessment procedures by exploring the contribution to be made by observation methods, with their greater flexibility and intrinsic validity; (3) in order to avoid the traditional pattern of comparing the contrasting programs according to a small number of critical criterion variables, it was decided to attempt to establish the areas of child functioning which each of them was most likely to strengthen, and to make differential predictions regarding the relative effectiveness of each form of program for producing specified sets of outcomes; (4) in light of the absence of reliable developmental norms, more extensive comparative studies of privileged and underprivileged children are needed in order to identify those areas of psychological functioning which are most impaired by disadvantaged status. Consequently, provision was made for the inclusion of a middle-class sample in the study in order to clarify the nature of social-class differences in development.

The major purpose of the study was to conduct a comparison of School Readiness and Child Development approaches to preschool education. Briefly characterized, the School Readiness method has as its prime goal the

preparation of children for effective functioning in their impending elementary school career. It focuses on the training of those concrete skills which relate most directly to the learning tasks the child is likely to encounter in kindergarten or first grade. The child is given explicit instruction in color identification, in the recognition and writing of numbers and letters, in clarity of diction, as well as in related facets of behavior which are presumed to facilitate an effective response to the demands of school life. This approach is pragmatic, concerned with fostering the explicit behaviors which appear to be most closely associated with school success. It is relatively atheoretical and is not directly concerned with aspects of psychological functioning which do not appear to have a close relation to effective school performance.

The Child Development approach more self-consciously addresses itself to the problem of psychological growth and integration. It attempts to create a learning environment and a set of relationships which permit mastery of materials and concepts, exploration and expression of feeling, and conflict resolution. Its goals then are broader; it is concerned more with cognitive growth and the child's overall developmental needs and less exclusively with preparation for school adjustment and academic success. It is more comprehensive but less focused, more individualized in its mode of operation and less didactic.

Sample Selection and Related Problems¹

The original data-gathering plan called for (1) the study of the educational environments created by the participating teachers, (2) the observational study of individual children, and (3) the pre-post test assessment of academic skills, conceptual functioning, curiosity, and degree of self-differentiation.

The field locations of the study were to be two classes from each of two Head Start centers that were exemplars of each of the two approaches to preschool education, and two classes from a nursery school for privileged children. Based on an estimate of 15 study children in each class, it was projected that the entire sample would consist of 60 children from four School Readiness Head Start classes, 60 children from four Child Development Head Start classes, and 30 middle-class nursery school children.

However, several major unforeseen problems forced a variety of profound modifications in design, all of which adversely affected the study. Most important, it was not possible to find suitable exemplars of either the Child Development or School Readiness approaches. In the light of the widespread publicity given to the Bereiter-Engelmann approach, which is avowedly a School Readiness orientation, it was assumed that many Head Start programs would be experimenting with this method. However, no such Head Start programs were in evidence in the New York City area. On the other hand, Child Development-oriented programs that were properly staffed and executed were

1. For a more detailed report of the material presented in this section, see Head Start Evaluation Program, Progress Report, March 1969.

almost equally rare. It should be noted that the search for such programs could begin only after Head Start classes were in session. Since in the final analysis teachers are responsible for the enactment of program, and centers were usually not able to specify in advance of the semester who would be teaching their class, let alone how the classes would be taught, observation of programs which were the most likely candidates for inclusion in the study could only proceed after the inevitable emergencies and organizational problems associated with the beginning of school had subsided and the teacher had had enough time with her class to feel that conditions had stabilized sufficiently for the classroom to be able to absorb visitors. At the same time, the pre-post test design of the study, and perhaps more urgent, the fact that the study was part of a national evaluation effort, narrowed the choice of the study samples and, by requiring that pretesting be completed within four to seven weeks after the class began, imposed sharp deadlines on the scheduling of data collection. As it was, it became impossible to comply with the specifications for the beginning of pretesting because of the complicated procedures involved in assessing and selecting prospective field locations and finally receiving the approval of the PAC of those centers with which we chose to work. Because of the long delays of several PAC groups to approve participation in our study and some last-minute complications which dictated the abandonment of one of our Child Development centers, the pretesting did not begin until mid-October in two of our sample centers and early November in the third.

In the final analysis, it was simply not possible to find the requisite study samples within the time and geographic confines of our work. Only one Child Development-oriented center -- the Ulysses Head Start Center² -- could be found whose director was both strongly committed to the child development approach to education and possessed a clear understanding of how to enact this approach, and whose PAC and staff were willing to cooperate. As previously noted, no Head Start center could be found with a curriculum plan, such as the one offered by Bereiter and Engelmann, whose singleminded concern is the improvement of the academic level of the children. While a number of Head Start center directors professed commitment to the academic training of their children, none seemed to have a systematic plan for accomplishing this goal. A not entirely satisfactory compromise led to the selection of the Clark Head Start Center, whose director and staff were trained in the Montessori method and whose program consisted of an amalgam of Montessori procedures and conventional nursery school methods. The Clark center was clearly not an ideal choice. While its focus was not derived from the current preoccupation with academic training for disadvantaged children, but rather from an effort to arrive at a new kind of early childhood education from a Montessori base, it was selected because it was committed to a definite instructional plan which departed significantly from the child development mode of education. The middle-class sample of children was selected from a child development-oriented nursery school, the Hillcrest School.

2. This name and those of the other sample centers are fictitious.

Because it was not possible to find suitable exemplars of the two educational approaches to be contrasted, and because only one center (instead of two) was found to provide the classes representing each approach, the number of children available for study was sharply reduced. In addition, we were faced with a number of constraining factors stemming largely from the specifications set forth by the national evaluation: only those classes could be used in which at least two-thirds of the children were new to Head Start, and only English-speaking children between $3\frac{1}{2}$ and $5\frac{1}{2}$ years old, with a maximum range of 12 months within classes, were eligible. On the basis of all these criteria, the number of children available was as follows:

Ulysses Head Start (Child Development): N = 49
Clark Head Start (School Readiness): N = 23
Hillcrest Nursery School: N = 18

By the time of pretesting, these already disastrously small N's had been further reduced by substantial attrition which continued throughout the year.³ The small size of the sample from each school group, combined with the overriding fact that these groups did not constitute satisfactory exemplars of the programs to be studied, fundamentally challenged the tenability of the study. Had it not been for the fact that our proposed study was embedded within our commitment to participate in the national evaluation study -- which entailed a large and complicated data-gathering process which overlapped with the procedures and goals of our comparative study -- it is unlikely that our study would have been undertaken under such adverse conditions. In effect, we decided to proceed with skeletal samples and inadequate prototypes of the programs to be studied with the hope that the ensuing study would serve to further clarify methodological and conceptual issues underlying this work, thus serving as a pilot effort for future investigations.

A final problem of major consequence stemmed directly from the study's involvement in the national evaluation of Head Start. The national evaluation called for the administration of a large battery of assessment measures -- of class, teachers, children and parents. The assessment load, in terms of time required for staff to be in the classroom, to test children and thereby to occupy limited testing space, was so extensive that the original data collection plan for the comparative study had to be seriously curtailed because testing space became unavailable and because there were limits to the frequency and concentration of intrusions that could be tolerated by the centers.

3. Attrition was mostly due (a) to children being withdrawn from the program, (b) to the dropping of children from the sample who had been designated as English-speaking by the teacher but turned out not to speak it well enough to be tested, and (c) to chronic absenteeism.

Description of Centers

Ulysses Head Start -- the Child Development Center

The Ulysses center was one of five centers sponsored by a church organization. It was located in a suburb just outside the central metropolitan area.

The houses in the immediate neighborhood of the center were two- or three-family dwellings in good condition, occupied in the main by black, low-income families. There were small businesses, a large park, and a public elementary school next door to the church that housed the Head Start classes. Because of overcrowding in this school, the church also accommodated one kindergarten and one first-grade class.

Approximately 60 children were enrolled in the four Head Start classes who with one exception were all Negro. There were two Head Start classrooms with two morning and two afternoon sessions. One of the classrooms was equipped primarily for arts and crafts; the other had large areas for block building, as well as a housekeeping corner and space for dramatic play.

At the beginning of the year there were two head teachers each teaching a morning and afternoon class, but in January one left and was replaced by her teacher aide, who became acting head teacher of the class. Both these teachers were Negro, between 40 and 45 years old. The head teacher had some college education, including courses in early childhood education. She had had several years' teaching experience with Head Start children, as well as preparatory training for Head Start, and had also taken a special training program in early childhood education. The acting head teacher had some high school education. She had had previous experience with Head Start and non-Head Start children and had also taken an eight-week preparatory course for Head Start.

During the year, a total of eight teacher aides came and went. Of the three who were in the two classrooms most consistently, two were Negro and one West Indian. Two were in their early forties and one between 22 and 27 years old. All three women said they had had some college education and two had taken courses in early childhood education. None had had any preparatory or special training. All had had previous experience with disadvantaged preschoolers, one with Head Start children.

The center director was involved in many types of interaction with the staff, children, and parents. There were formal conferences on each child, daily informal gatherings, and numerous parent committee meetings and staff meetings. Though the director's relations with her staff and parents were positive in many respects, there was friction. This seems in part to have been due to the director's strong desire to develop and implement her educational ideas, on the one hand, and her frustration at her staff's lack of training and enthusiasm, on the other.

PS 003671

The Program. One of the striking aspects of the program at this center was the joining at certain times of the Head Start classes with the public school kindergarten and first-grade children who were housed in the same building but belonged to the public school next door. At these times, there was a free flow of children from all these classes through all rooms, dictated only by choice of activity, e.g., a child might elect to do collage during free play, in which case he would go to the classroom housing the arts and crafts equipment. The center director was able to put this innovation into practice partly because the Head Start, kindergarten and first-grade teachers all had the same child development orientation (thus also providing program continuity to the second grade).

The two head teachers in the Head Start classrooms were markedly different. One was open, soft-spoken, calm, and firm in her handling of the children. She actively helped them become more aware of their environment -- perhaps bringing in a bee in a jar to observe and discuss, or having each child comment on the sounds he could hear from the outside. She often used play activities to teach cognitive skills, e.g., when a child was painting she would point out colors and color contrasts. Her room was generally full of children from various classrooms working busily but moving freely from one activity to another. Although it was often noisy, this teacher did not suppress noise unless it was disruptive.

The other teacher, who had been the teacher aide, was a woman of strong personality who seemed to have exerted a greater influence over the children than the head teacher whom she had replaced. Her teaching approach was individualized but she had excellent rapport with the class group. Although her manner sometimes seemed to overwhelm the children, they clearly responded to her interest in them and her general enthusiasm and cheerfulness.

Clark Head Start -- the School Readiness Center

Until 1968, Clark was sponsored by a large community service organization. Early in 1968, however, the center incorporated independently. Funded by CAP, it continued to operate only the one center.

Clark was located in the middle of an inner city slum. As one walked the streets, the picture was grim. Rusted, dismembered cars, garbage, boarded-up buildings, rats, drunks, and addicts were all in evidence. The ethnicity of the area, as reflected in the child populations of its two public schools, was roughly 75% Puerto Rican, 22% Negro, and 3% other white. The principals of the two public schools who gave us these figures also told us that the mobility in the area was almost 100%; newly arrived Puerto Rican families settled only long enough to find jobs and better housing elsewhere.

The church which housed the Head Start classrooms was a bright spot amidst the desolation, set in a small patch of trees and grass. The building was old, but solid and well kept. There were four classes with an enrollment of something over 60 children, approximately half of whom were Spanish speaking, Puerto Rican, and half Negro. The classrooms were housed in a semi-basement.

The testing space allocated to us left much to be desired. Located below the classroom level, children had to be taken down some stairs, along a dimly-lit corridor, and the examiner had to unlock two doors in order to reach an ill-ventilated room with small windows at ceiling level. The children, however, appeared in no way disconcerted by these surroundings.

Due to a variety of complex circumstances, three of the four sample classes had a new teacher in the middle of the year, one of whom was new to the center. The two teacher aides remained throughout the year.

The children, too, were constantly being transferred from class to class for reasons often unclear to us. In addition, because of the high mobility of the neighborhood population, children were regularly withdrawn from the program and new ones enrolled. It is difficult to know what impact all these changes had upon the children; a degree of confusion was evident at all levels.

The three head teachers were all white and relatively young women between 22 and 33 years old. All had baccalaureate degrees and had taken courses for credit beyond the degree, as well as graduate courses in early childhood education. In addition, they had all had between 12 and 18 months' Montessori training. None had had prior training for Head Start and only one had had previous teaching experience with Head Start children, though the other two had taught non-Head Start disadvantaged preschool children.

Of the two teacher aides, one was a Negro, between 34 and 39 years old, with some high school education, and the other was Puerto Rican, a high school graduate in her mid-twenties. Both had had previous experience with disadvantaged preschool children -- Head Start and non-Head Start -- and both had had Head Start in-service training.

The atmosphere of the center as a whole was pleasant and congenial if a bit lacking in organization. The director and her staff related well to one another as well as to the children and parents. The latter were quite active but their activities were social rather than program-related. There were coffee hours, sewing clubs, and dances but comparatively little time was given to parent-teacher conference-type activities.

The Program. Clark classified its program as "modified" Montessori. While not embracing all the orthodoxies of the Montessori method, the program operation relied heavily on Montessori materials as working tools for the achievement of intellectual competence in pre-reading, writing and mathematical concepts. The teachers encouraged individual work and geared their own activities to individual instruction and supervision in the use of the Montessori equipment. Group work was infrequent and social interaction among the children discouraged.

Two of the teachers represented the clearest exemplars of the Montessori method. They were involved almost totally with structured work situations which were geared to individual children and centered on the Montessori equipment. Though they did not deny the affective life of the child, these two teachers considered their main job to be one of cultivating the intellect, and their lack of emphasis on or attention to the social development

of the children was striking. They maintained a firm, no-nonsense control over the children.

When these teachers were asked, in the teacher interview, to choose the four concepts (from a list of 24) which best described their teaching approach, both included "transmitter of knowledge and skills" and "preparer of an environment which is conducive to independent learning" among the four. Similarly, when asked what the children did that pleased them, both teachers said they were pleased when the children were happy in their work and gaining a sense of achievement. When asked what they would do and say when the children did something that pleased them, both these teachers replied that they would say or do nothing. This orientation, as mentioned earlier, was consistent for all three teachers.

The third head teacher presented a somewhat different picture. She was much less of a disciplinarian, more spontaneous, and kept less distance between herself and the children in her role as an adult authority. She was the only one of these three teachers who sat down on the floor with the children in order to enter wholly into their activities. She too tended to concentrate on individual children, but in a personal rather than work-related way. She was often so involved with a single child that the rest of the class was neglected. The result was an atmosphere of confusion and chaos which was compounded by a basic lack of structure in her classroom. This teacher relied much less than the others on the Montessori equipment but did not seem to be implementing an organized program. When asked to choose the four concepts (out of 24) which best described her teaching approach, she said she considered herself first of all as a person, not a teacher.

Hillcrest Nursery School

Hillcrest Nursery School was located in a large, sprawling housing project in a residential section of the city. The apartment units of the project were connected by winding pathways bordered by many bushes, trees and grassy spaces. There were small businesses, shops, cooperative apartments and some semi-detached houses nearby, all of which combined to create a typical middle-class setting. The population of the neighborhood, as reflected in the school, was primarily Irish, Jewish, and Italian, with probably less than 2% Oriental and Negro.

The school occupied a one-story, T-shaped building in the project and had its own yard. Now in its fifteenth year, it was a fully integrated community nursery school composed of eight classes in double session, with a total enrollment of approximately 160 children, the majority of whom were white. The school had an elected Board of Directors made up of parents, as well as an Advisory Board which included professionals.

The parents paid to send their children to Hillcrest and the mothers were required to give time either as "mothers' helpers" in the classroom or by serving on committees. Scholarships and partial scholarships were available. The parent group was young, mobile and middle class, and most of the fathers and some mothers were business or professional people.

The majority of the children did not come from the housing project. The reasons for this were threefold: the parents in the development could not afford it, they did not have the time to contribute the activities required of parents in the program, and many of them worked and needed day care, rather than a three-hour program for their children.

The director was an extremely competent woman. She felt her staff met the standards she had set for herself: a commitment to the importance of education for young children and a respect for each child as an individual. She offered her staff a unique in-service training program which she was well-equipped to do, having served as a field consultant to a university Head Start in-service training program. She felt that parents benefited from having an active role in the school. By working in the classroom they could gain insight and awareness into the kinds of experiences their children were having, and through committee work they assumed responsibilities necessary to the appreciation of what a good preschool can provide. These experiences also acted as a springboard for the parent-teacher conferences.

The Program. One morning and one afternoon class were selected for the sample in this school. Both classes shared the same room and equipment. In both there was a varied indoor program, stressing arts and crafts, and a long, active outdoor play period. The children tended to gravitate more toward activities which involved a group rather than solitary activities. The group social life was the most conspicuous feature of both classes. The teachers were active in meeting the children's needs, but extended one-to-one sessions with a child were rare.

The teacher of the morning class was effusive with the children, even overwhelming, sometimes injecting a note of tension. Her mode of teaching tended to be in the direction of encouraging the children to extend or elaborate their play activity. The children were highly sociable and independent and required little help or attention from the teacher. They formed groups of five or six, often even larger, for extended episodes of dramatic play, on a wide range of themes. Nearly every child in the class participated in this kind of group play.

The teacher of the afternoon group was more restrained, and somewhat didactic. She spent a lot of time trying to draw more complex meanings or additional facts from play experiences. In particular, she seemed to push for self-reliance, often parrying a question or request with a suggestion that the child could help himself. This class, on the whole, was quieter than the morning class. The afternoon children were also sociable, but less so than the morning children, perhaps because they were some months younger. However, there was a more or less permanent group consisting of the more mature children, whose members played together most of the time.

The Instruments of the Study

The instruments that were used to study these groups were a conglomerate of measures, in part prescribed by the national evaluation study and in part selected as being especially relevant to the purposes of the comparative study. It should be noted, therefore, that the choice of instruments was not exclusively determined by the needs of the comparative study. In

this report, only those instruments prescribed by the national evaluation which are relevant to the study will be included. A list of the instruments of the study is presented below, together with brief descriptions where appropriate.⁴ All except the individual child observations were given on a pre-post basis.

Stanford-Binet*

Inventory of Factors
Affecting Stanford-Binet
Test Performance*

Preschool Inventory*

Animal House, WPPSI*

Matrix Test[#]:

A test for assessing classification, sorting and related cognitive skills associated with inferential reasoning.

Kaleidoscope[#]:

A measure of curiosity and exploration.

Figure Drawing:

A self-drawing.

Individual Child
Observations[#]:

A continuous record of the child's activities and behavior during two observation cycles, each consisting of two 15-minute observations made on different days and spanning various activities.

Observation of Spon-
taneous Language[#]:

A measure to categorize the spontaneous, interindividual communications of preschool children in class.

The sample N's for each measure are presented below. Only those children who were available for pre- and posttesting are included.⁵

4. Instruments with an asterisk were prescribed by the national evaluation study. For a discussion of these techniques, see Head Start Evaluation Study, Progress Report, March 1969.

Instruments with a # are presented in greater detail in the Appendix.

5. The ethnic composition of this sample was as follows:

Ulysses: all Negro

Clark: all Negro

Hillcrest: 83% white, 17% Negro, Oriental, and mixed.

Distribution of the Sample for Each Measure

	<u>Ulysses</u>	<u>Clark</u>	<u>Hillcrest</u>
Stanford-Binet (Factors Affecting Stanford-Binet Performance)	39	17	15
Preschool Inventory	38	15	14
Animal House, WPPSI	38	15	14
Matrix Test	19	10	12
Kaleidoscope	20	12	12
Figure Drawing	22	10	12
Individual Child Observations ⁶	12	12	12
Observation of Spontaneous Language ⁶	10	10	5

Results

While the focus of this project was the comparative study of the influence of different types of Head Start programs, it was also designed to provide a comparison of the functioning of a sample of disadvantaged preschool children with an advantaged group both at the outset and the conclusion of a year of preschool experience. Thus the findings will be examined in terms of differences between groups as a function of varying social-class background as well as in terms of the nature of their preschool experience.

The ages of the children at the beginning of pretesting are given in Figure 1. It may be seen that the two Head Start groups were essentially equivalent in mean age and that the middle-class Hillcrest group was somewhat younger. The range of ages was somewhat greater in the Clark group than the twelve-month span in the Ulysses and Hillcrest groups.

Stanford-Binet

Before examining the Stanford-Binet data, it is important to note certain features of the method of test administration which have implications for the interpretation of results. While there was some overlap in examiners for the Ulysses and Clark groups, the center groups, in large measure, were tested by different examiners in the pretest. Further, the post-testing was almost always done by a different examiner from the one who had done the pretest for a given center. Thus, differences between groups, and differences between pre- and posttest scores within as well as between groups, are contaminated with examiner differences. The addition of the Birch work-nonwork recording system of test performance lengthened the time of test administration by ten to thirty minutes, so that each test-taking session was longer than is typically found with the Stanford-Binet. In addition, the Birch procedure demanded that all subparts of all items be administered regardless of success or failure, thus greatly increasing the child's fatigue and probably emphasizing a sense of failure for some children.

6. The Head Start sample on these measures was matched for age, sex, ethnicity and IQ.

Figure 1

Distribution of Ages at Time of Pretesting

Stanford-Binet pretest

<u>AGE IN MONTHS</u>	<u>Ulysses (N = 39)</u>	<u>Clark (N = 17)</u>	<u>Hillcrest (N = 15)</u>
62			
61			
60	X		
59			
58	XXXXXXX	X	
57	XXX	X	
56	XX	X	XX
55	X	XXX	
54	XXXX	X	
53	XXX	XXXX	X
52	XXXX	X	
51	XX		XX
50	XXXX		X
49	XX	XX	XXXX
48	XXXX	X	
47	XX		XX
46			XXX
45		X	
44			
43			
42		X	
	$\bar{X} = 53.2$	$\bar{X} = 52.2$	$\bar{X} = 49.7$

PSI pretest

<u>AGE IN MONTHS</u>	<u>Ulysses (N = 39)</u>	<u>Clark (N = 15)</u>	<u>Hillcrest (N = 14)</u>
62	X		
61	XXVXXXX	X	
60	XX		
59	X		
58	X	XX	
57	XXXX	XXX	X
56	X	XX	X
55	XXX	XXX	
54	XXXX		X
53	XX		X
52	XXXX		X
51	XXXX		X
50	XXXX	XX	XX
49	X		X
48			X
47		X	XXXX
46		X	
45			
44			
43			
42			
	$\bar{X} = 55.3$	$\bar{X} = 54.5$	$\bar{X} = 50.6$

The distribution of pretest Stanford-Binet IQ scores for each school group is given in Figure 2. It may be observed that the means for the two Head Start center groups were close to each other and only slightly below the general population mean IQ of 100. It would appear from these data that the children at these centers were functioning at a higher level than those usually encountered among Head Start samples. The mean IQ of the advantaged group was approximately one standard deviation above that of the Head Start children.

Before turning to the changes in IQ scores that were found in the post-testing, it is useful to examine how pretest performance varied from item to item. Table 1 presents the proportion of Ss in each preschool group that passed each Stanford-Binet item on the pretest administration. It may be observed that performance on items below the four-year level was almost uniformly high for all three groups. The only striking exception to this pattern was the substantially lower level of performance of the Ulysses group on the Picture Vocabulary test (III-2) and the even lower performance by the Clark group on the Comprehension item (III-6,6) which called for responses to the questions: "What must you do when you are thirsty?" and "Why do we have stoves?" Most of the children in all three groups were able to string beads, build a bridge, show memory for pictures, copy a circle, draw a vertical line, compare the size of balls, fit two cards together to complete a figure, discriminate among pictures of animals, tell about a picture, and sort buttons into black and white groups. While performance was high on these items for all three groups, the proportions of children answering them correctly were usually highest in the Hillcrest school.

Performance on the more difficult items showed more discrepancy among the groups. At the four-year level, the Clark children performed least well by a large margin. They rarely succeeded on the Picture Vocabulary and Comprehension items ("Why do we have houses?" "Why do we have books?") and ranked last by a considerable margin on Opposite Analogies, and on a Pictorial Identification item ("Show me the one that we cook on"). A reversal of this pattern occurred on the item which required naming objects from memory; here Hillcrest children performed least well while the item was passed by all the Clark children. The Clark children continued to perform least well at the four-and-a-half-year level, scoring lowest on all six items at this level. They showed special difficulty with the Opposite Analogies item, with questions which ask what materials familiar objects are made of, and with an item requiring them to perform three commissions. The Ulysses children, too, were unable to cope successfully with the Opposite Analogies and the What-materials-are-familiar-objects-made-of items. Items calling for aesthetic comparisons, pictorial similarities and differences, and comprehension (i.e., "What do we do with our eyes?" "...with our ears?") were passed by most of the children in all three groups.

Curiously, as the items increased in difficulty -- at the fifth-year level -- it was the Clark children's performance which more closely resembled the more competent levels achieved by the Hillcrest group, and the Ulysses group whose performance lagged behind the others. The Ulysses group performed considerably worse on the paper folding and pictorial similarities and differences items -- which were difficult for all the groups -- and were

Figure 2

Distribution of Stanford-Binet Pretest IQ Scores

<u>IQ Score</u>	<u>Ulysses (N = 39)</u>	<u>Clark (N = 17)</u>	<u>Hillcrest (N = 15)</u>
140-144			X
135-139			
130-134			
125-129	X		X
120-124	X	X	XX
115-119	XXX		XXXXX
110-114	XX	X	XX
105-109	XXXXXXXXXXXX	XX	
100-104	XXXXXX	XX	
95-99	XXXXX	XXXX	X
90-94	XXX	XXX	XX
85-89	X	X	
80-84	X	X	
75-79			X
70-74	XX	XX	
65-69	X		
60-64	X		
55-59			
50-54	X		
	$\bar{X} = 99.4$ $s = 16.42$	$\bar{X} = 96.4$ $s = 12.23$	$\bar{X} = 111.9$ $s = 15.55$

Table 1

Percent of Ss Answering Stanford-Binet Items Correctly on Pretest
(Age Level Items III-VI)

Age Level	Item	Item Description	Ulysses (N = 39)	Clark (N = 17)	Hillcrest (N = 15)
III	1	stringing beads	100	100	100
	2	picture vocabulary	67	94	100
	3	block building: bridge	97	100	100
	4	picture memories	100	100	93
	5	copying a circle	95	94	93
	6	drawing a vertical line	95	100	100
III-6	1	composition of balls: which ball is bigger?	87	88	87
	2	patience: pictures -- putting 2 cards together	79	82	100
	3	discrimination of animal pictures	92	88	100
	4	response to pictures: "Tell me about it"	87	94	100
	5	sorting buttons (black and white)	92	88	93
	6	comprehension: what do we do when thirsty? why do we have stoves?	85	59	74
IV	1	picture vocabulary	54	35	67
	2	naming objects from memory	87	100	67
	3	opposite analogies	79	53	93
	4	pictorial identification: show me what we cook on show me what we carry when it rains	82	65	93
	5	discrimination of forms	92	82	87
	6	comprehension: why do we have houses? why do we have books?	72	24	67
IV-6	1	aesthetic comparison	87	76	67
	2	opposite analogies	38	24	80
	3	pictorial similarities and differences	74	65	80
	4	materials: what is a house made of? what is a window, book, made of?	36	7	67
	5	three commissions	64	35	74
	6	comprehension: what do we do with eyes? what do we do with ears?	69	59	74
V	1	picture completion: man	26	13	40
	2	paper folding: triangle	10	30	33
	3	definitions: what is a ball, hat, stove?	44	71	80
	4	copying a square	10	13	20
	5	pictorial similarities and differences	26	59	67
	6	patience: rectangles ▽△	13	20	20
VI	1	vocabulary	18	0	13
	2	differences: bird and dog, slipper and foot, wood and glass	16	7	33
	3	mutilated pictures	10	7	20
	4	number concepts: give me ____ blocks	3	0	13
	5	opposite analogies	8	0	26
	6	maze tracing	13	20	33

also considerably less effective than the other two groups on the easiest item of this level -- one which asked for definitions of a set of common objects. None of the groups showed much ability to copy a square or to fit two triangles together to form a rectangle. Only the Hillcrest group showed relatively substantial ability in completing the figure of a man.

At the six-year-old level, only the Hillcrest group showed even a slight ability to cope with items which called for providing differences between objects sharing certain attributes, e.g., a slipper and a boot, wood and glass, and with the opposite analogies, and maze tracing items. The vocabulary items of this level, a mutilated pictures item, and an item calling for cardinal numerical thinking were almost universally failed by all the groups.

The change in IQ scores from pre- to posttesting are presented in Figure 3. The interval between pre- and posttesting was generally about six months; pretesting occurred at a point from five to seven weeks after the classes had begun and the pre-post test interval varied from six to seven months at Clark and Hillcrest and from four and a half to seven months at Ulysses.⁷ The change score data indicate widespread shifts in test performance, i.e., 49 percent of the Ulysses group, 47 percent of the Clark group and 73 percent of the Hillcrest group changed IQ by at least five points. What is of special interest, however, is that only in the Hillcrest school was there actually a rise in mean IQ, and a substantial rise at that.

There are three bases for change in IQ score from pre- to posttesting: (1) performance declines, i.e., items are failed in the posttest that were passed in the pretest, (2) performance is essentially unchanged, i.e., the same mental age is maintained while chronological age rises, and (3) performance improves, i.e., the child passes items in the posttest that he failed in the pretest, above and beyond what might be expected because of his increase in chronological age.

Examination of pre-post differences in item performance (Table 2) indicates that almost all items below the four-year level that were failed during pretesting were passed during posttesting. The two striking exceptions to this pattern were both found in the Ulysses group. A quarter of their sample was still not able to pass the third-year Picture Vocabulary item, and unlike the other two groups, those children who had previously failed the Comprehension item at the III-6 level continued to fail this item at posttesting. At the four-year level, it is not surprising to find more evidence of improvement, since these items had been failed more widely than those at preceding levels at pretest, and were more likely to be passed more often at the posttest. The most growth across all groups occurred in the Picture Vocabulary item although sizeable gains were made in other fourth-year items as well. The greatest improvement at the four-year level was shown by the Clark group, especially in the Comprehension problem, the Picture Vocabulary item and the Opposite Analogies item.

7. The short minimum pre-post interval at Ulysses was due to delayed pre-testing because this center was the last to open and the PAC the last to approve the study.

Figure 3

Distribution of Stanford-Binet Change Scores

<u>CHANGE SCORE</u>	<u>Ulysses (N = 39)</u>	<u>Clark (N = 17)</u>	<u>Hillcrest (N = 15)</u>
+ 30-34			X
+ 25-29			XX
+ 20-24			
+ 15-19	X		X
+ 10-14	XX	X	XXX
+ 5-9	XXXX	XXX	XXX
+ 1-4	XXXXXX	XXXX	XXX
0	XXXX	X	
- 1-4	XXXXXXXXXX	XXXX	X
- 5-9	XXXXXX	X	
- 10-14	XXX	X	
- 15-19	XX	X	X
- 20-24			
- 25-29		X	
- 30-34	X		
Mean Pretest IQ	99.4	96.4	111.9
Mean Post-test IQ	97.1	94.7	121.9
Mean Change Score	-2.3	-1.7	+10.0

U = Ulysses (N = 39)
C = Clark (N = 17)
H = Hillcrest (N = 15)

Table 2

- 135 -

Pre-Post Performance on Individual
Stanford-Binet Items (Age Level III-VI)

Items	++			+-			-+			--		
	U	C	H	U	C	H	U	C	H	U	C	H
III												
1 stringing beads	39	17	15	-	-	-	-	-	-	-	-	-
2 picture vocabulary	26	16	15	-	-	-	3	1	-	10	-	-
3 block building: bridge	38	17	15	-	-	-	-	-	-	1	-	-
4 picture memories	39	17	14	-	-	-	-	-	1	-	-	-
5 copying a circle	37	16	14	-	-	-	2	1	1	-	-	-
6 drawing a vertical line	36	17	15	1	-	-	1	-	-	1	-	-
III-6												
1 composition of balls: which one is bigger?	33	15	12	1	-	1	5	2	2	-	-	-
2 patience: pictures -- fitting 2 cards together	30	14	14	1	-	1	6	3	-	2	-	-
3 discrimination of animal pictures	36	15	15	-	-	-	2	2	-	1	-	-
4 response to pictures: "Tell me all about it"	34	16	15	-	-	-	3	1	-	2	-	-
5 sorting buttons (black and white)	36	15	14	-	-	-	1	1	1	2	1	-
6 comprehension: what do when thirsty? why do we have stoves?	32	10	11	1	1	-	-	5	4	6	1	-
IV												
1 picture vocabulary	18	6	8	3	-	2	9	7	4	9	4	1
2 naming objects from memory	33	17	9	1	-	1	4	-	5	1	-	-
3 opposite analogies	27	9	14	4	-	-	3	5	1	5	3	-
4 pictorial identification: show me what we cook on; show me what we carry when it rains	31	11	14	1	1	-	3	1	1	4	-	-
5 discrimination of forms	34	14	13	2	-	-	2	3	2	1	-	-
6 comprehension: why do we have houses? why do we have books?	26	4	10	2	1	-	4	9	4	7	2	1
IV-6												
1 aesthetic comparison	31	12	9	3	1	1	2	3	5	3	1	-
2 opposite analogies	14	3	11	1	1	1	9	4	3	15	8	-
3 pictorial similarities and differences	26	9	12	3	2	-	6	5	2	4	1	1
4 materials: what is a house made of? what is a window, book, made of?	12	1	10	2	-	-	13	2	3	12	14	2
5 three commissions	16	6	8	9	-	3	9	5	3	5	4	1
6 comprehension: what do we do with eyes? what do we do with ears?	24	7	11	3	3	-	6	2	3	6	5	1
V												
1 picture completion: man	7	2	6	3	-	-	5	5	8	22	10	1
2 paper folding: triangle	2	4	5	2	1	-	9	3	6	26	9	4
3 definitions: what is a ball, hat, stove?	12	12	11	5	-	1	3	4	1	19	1	2
4 copying a square	4	2	3	-	-	-	9	8	6	26	7	6
5 pictorial similarities and differences	10	10	9	-	-	1	7	4	4	22	3	1
6 patience: rectangles	2	1	2	3	2	1	11	2	8	23	12	4
VI												
1 vocabulary	2	-	2	5	-	-	6	2	9	26	15	4
2 differences: bird and dog, slipper and foot, wood and glass	2	1	4	4	-	1	3	2	2	30	12	8
3 mutilated pictures	4	-	2	-	1	1	6	3	3	29	12	9
4 number concepts: give me _____ blocks	1	-	1	-	-	1	4	3	3	34	13	10
5 opposite analogies	2	-	4	1	-	-	3	1	8	33	15	2
6 maze tracing	2	2	4	3	1	1	10	4	7	24	9	3

Failure on both the pre- and posttest was relatively rare at the four-year level. On the six items at the four-year level, there were a total of 27 instances of failure on both the pre- and posttesting among the 39 children of Ulysses, nine such instances from among the 17 children from Clark, and only two such cases among the 15 Hillcrest children.

Equally large gains were found at the four-and-a-half-year level in all three groups. The largest gains were made in the Opposite Analogies item, in the ability to tell what materials familiar objects are made of and the ability to perform commissions. The Ulysses group showed a remarkable spurt in the Materials item. Whereas there were 45 instances of failure on both pre- and posttesting of this cluster of six items among the Ulysses group, and 33 in the Clark group, there were only five such instances in the Hillcrest sample. In particular, the Opposite Analogies and Materials items presented widespread difficulty for the Ulysses and Clark groups (the Clark group had special difficulty with the Materials item) but not for the Hillcrest sample.

Large gains were also recorded at virtually all the five-year items -- Picture Completion, Paper Folding, Copying a Square, Pictorial Similarities and Differences, and the fitting together of two triangles to form a rectangle. Especially noteworthy was the major improvement in completing the drawing of a man in the Hillcrest group, an item in which this group was already advanced in the pretest performance, and also, the vast improvement in fitting triangles to form a rectangle in the Ulysses and Hillcrest groups. Five of the six items at the five-year level were failed on both the pre- and posttest by more than half the children from Ulysses. In all, at this level, there was a total of 138 failures at both the pre- and posttesting in the Ulysses group, compared with 42 at Clark and 18 at Hillcrest.

The greatest amount of progress at the six-year level was made in the Maze Tracing task. Hillcrest showed very great improvement in the Vocabulary and Opposite Analogies items.

There were, in addition, a surprising number of instances in which children who passed an item on the pretest failed the same item on the posttest (see Table 2). While these retrogressions were for the most part distributed throughout the range of items, they occurred more often in some items than others. Retrogressions occurred for more than three of the total of 71 children (an arbitrary cut-off point) in the study on the following items: Picture Vocabulary (IV), Opposite Analogies (IV), Aesthetic Comparisons (IV-6), Pictorial Similarities and Differences (IV-6), Three Commissions (IV-6), Comprehension (IV-6), Definitions (V), Patience (fitting triangles to form a rectangle) (V), Vocabulary (VI), Differences (VI), and Maze Tracing (VI). It is not surprising to find instability of response to items which call for a restructuring of the perceptual field since such items are most susceptible to lapses in attention and variations in frames of reference. On the other hand, retrogressive changes in the ability to define words (Vocabulary) are more difficult to explain except in terms of variations in the child's attentiveness to the test and/or differences in the examiners' judgments in scoring. Perhaps most surprising among these shifts backward

was the considerable retrogression in the Three Commissions item (IV-6) in the Ulysses group: nine of the 39 Ulysses children failed this item on the posttest after passing it initially. For the test as a whole, most of these retrogressions occurred in the Ulysses group. Compared with a total of 64 instances in Ulysses, there were 15 in Clark and 17 in Hillcrest.

Inventory of Factors Affecting Stanford-Binet Test Performance

The rating of Stanford-Binet pretest performance would appear to suggest that in all three groups the majority of the children's performance was not adversely affected (see Table 3). For the most part, differences among the three school groups were small. The largest differences among groups were with regard to the acceptability of tasks -- i.e., more of the Hillcrest children seemed to be hostile to the pretest; the degree of persistence that was shown -- Clark children were rated as being somewhat less persistent on the pretest; the greater social ease found among Ulysses children; the lesser responsiveness to encouragement among Hillcrest children; and the markedly more impaired patterns of activity level and verbal expression in the Clark children.

Differences among the groups in their change from pre- to posttesting were more dramatic. In contrast with the tendency for Hillcrest children to perform with less impairment during posttesting and the relatively even level maintained by the Ulysses group from pre- to posttesting, the Clark group's performance was judged to be markedly more impaired during posttesting. It is important to note, however, that a single examiner did all the posttesting in this group and his ratings may have been idiosyncratic.

The Hillcrest group was found to perform considerably better at posttesting in relation to its response time, i.e., fewer children tended to give up easily, in the acceptability of the tasks, in their degree of social ease, their activity level, i.e., fewer children were either hyperactive or hypoactive. In the Ulysses group, factors affecting test performance did not change markedly from pre- to posttesting. Exceptions to this pattern were changes in response time, i.e., many fewer children needed urging to respond, and also declines which were recorded in their persistence and activity patterns. Ratings of the Clark children declined on every dimension except activity level. They were seen to be markedly less attentive, less realistic in their sense of competence, slower to respond, less accepting of the test tasks, much less persistent, less realistic in their reaction to failure, less at ease socially, and less responsive to normal encouragement. It is interesting that all three school groups were rated as declining in persistence from the pretesting to the posttesting.

Table 3

Percent of Children Who Were NOT Adversely Affected
by Stanford-Binet Pre- and Posttesting

	Ulysses (N = 39)		Clark (N = 17)		Hillcrest (N = 15)	
	Pre	Post	Pre	Post	Pre	Post
1. Gives test attention	72	64	75	50	67	73
2. Realistic sense of competence	74	77	81	56	73	67
3. Adequate response time	62	87	56	31	60	93
4. Accepts task	87	79	88	63	67	87
5. Adequately persistent	85	72	69	31	80	67
6. Reacts to failure realistically	85	92	88	50	73	80
7. Socially at ease	90	85	69	44	73	93
8. Responds to normal amount of encouragement	85	90	88	38	67	80
9. Normal activity level	82	67	50	56	73	100
10. Normal verbal expression	51	49	25	19	67	73

Preschool Inventory (PSI)

The mean pretest raw scores (see Table 4) present a somewhat different pattern from that obtained with the Stanford-Binet IQ pretest data. While the rank order of the school groups is the same as that for the Stanford-Binet, the mean pretest score of the Ulysses group is much closer to the Hillcrest children than to the Clark children. This different pattern may be largely attributable to the fact that the IQ score was able to take into account the lower age of the Hillcrest group.

Table 4

Mean PSI Scores for Pre-Post Testing

	<u>Ulysses</u> <u>(N = 38)</u>	<u>Clark</u> <u>(N = 15)</u>	<u>Hillcrest</u> <u>(N = 14)</u>
Pretest	35.8	30.2	37.6
Posttest	<u>40.2</u>	<u>39.4</u>	<u>47.6</u>
Change	4.4	9.2	10.0

An examination of performance on each item of the pretest indicates, curiously, that many more Hillcrest children failed to give their first name when it was asked of them but were able to tell their age (see Table 5). Most of the children were able to follow simple directions but, in marked contrast with the other groups, no children from Clark apparently knew the meaning of the word "wiggle." The Clark group also lagged considerably in the ability to follow complex directions involving color, e.g., put red car on the black box; the Hillcrest children usually scored highest on these items. The Clark children also lagged substantially on items calling for general information. The few items which called for the simulation of motion made by familiar objects were failed by most of the children in all the groups. Items which asked the number of elements in familiar objects produced variable results, with those aggregates involving larger numbers failed more often. The Hillcrest children were somewhat better in making comparative judgments of objects of unequal size or number or some other attribute. All the groups were able to match figures with familiar objects which they schematically resembled, but were largely unable to copy simple figures. The Hillcrest group showed the greatest ability to name colors and the Clark group was least competent in this area.

The rise in mean score from pretest to posttest was twice as great in Clark and Hillcrest as it was in Ulysses. This was mostly due to the greater gains in performance which occurred in these groups, but also to the fact that the Ulysses group showed many more instances of decline in performance from pretest to posttest. The Hillcrest children's rise in score occurred most sharply on the item calling for them to give their first name and in commissions involving complex relational terms such as under and second. The Clark children showed especially large gains on commissions which required distinguishing among colors, and those items which called for information about what a teacher, dentist, etc., does.

U = Ulysses (N = 38)

C = Clark (N = 15)

H = Hillcrest (N = 14)

- 16a -

Table 5

Pre-Post Performance on Individual PST Items

Items	++			+-			-+			--		
	U	C	H	U	C	H	U	C	H	U	C	H
1. What is your first name?	32	14	9	2	0	0	3	1	4	1	0	1
2. How old are you?	20	8	11	4	1	1	9	4	2	6	2	0
3. What is your last name?	2	2	1	2	1	0	10	4	7	25	8	6
4. Show me your shoulder.	25	12	8	4	1	2	8	2	3	2	0	1
5. Show me your heel.	18	4	3	8	3	0	5	1	2	8	7	9
6. What call (finger)?	21	9	11	3	1	1	7	5	1	8	0	1
7. What call (knee)?	23	7	10	4	3	0	8	3	2	4	2	2
8. What call (elbow)?	15	6	7	2	2	0	14	1	6	8	6	1
9. Raise your hand.	38	14	13	0	0	0	0	1	1	0	0	0
10. Jump.	38	13	14	1	0	0	0	2	0	0	0	0
11. Hello very loudly.	27	9	9	5	1	0	4	4	4	3	1	1
12. Wiggle.	17	0	8	11	0	2	5	3	3	6	12	1
13. 3 cars in big box.	15	7	8	6	0	1	12	5	4	6	3	1
14. Red car on black box.	19	3	9	6	0	1	5	8	2	9	4	2
15. Yellow car on little box.	7	1	3	10	2	1	7	6	2	15	6	8
16. Blue car under green box.	14	3	8	5	0	0	4	4	5	16	8	1
17. 2 cars behind middle box.	8	3	1	8	0	0	4	5	9	19	7	4
18. Give everything to me.	19	7	10	7	0	1	8	8	3	5	0	0
19. Who go to if sick?	22	8	8	5	3	1	5	3	1	7	1	4
20. Where find boat?	21	6	9	7	1	1	5	2	4	6	6	0
21. Where buy gas?	24	5	9	5	1	1	5	1	4	5	8	0
22. When breakfast?	15	3	4	7	2	3	7	6	5	10	4	2
23. What do to read something?	23	4	5	4	2	2	7	6	6	5	3	1
24. Where find lion?	14	6	8	3	1	1	11	3	4	11	5	1
25. What does mother do?	24	7	8	7	1	2	5	7	3	3	0	1
26. What does dentist do?	18	6	5	5	1	2	9	7	4	7	1	3
27. What does teacher do?	9	0	2	4	2	2	11	8	6	15	5	4
28. Which way water fall?	2	1	3	5	2	1	8	1	3	24	11	7
29. Which way record?	10	5	6	11	1	0	7	5	5	11	4	3
30. Which way ferris wheel?	4	0	1	11	2	1	5	4	5	19	9	7
31. How many eyes?	31	11	13	3	1	1	3	2	0	2	1	0
32. Count (to 5).	23	11	9	5	1	2	5	3	3	6	0	0
33. How many hands?	22	5	8	2	2	1	8	4	0	7	4	5
34. How many wheels - bicycle?	11	4	5	8	2	0	11	6	3	9	3	6
35. How many wheels - car?	6	4	5	8	3	1	12	1	5	13	7	3
36. How many wheels - tricycle?	3	3	4	4	2	0	10	2	3	22	8	7
37. How many corners - paper?	6	0	5	3	0	4	10	2	0	20	13	5
38. How many toes?	0	0	1	0	0	0	2	1	3	37	14	10
39. Bigger - tree or flower?	30	9	11	2	1	1	5	5	2	2	0	0
40. Slower - car or bicycle?	19	4	5	11	3	2	6	4	5	3	4	2
41. Heavier - brick or shoe?	21	3	10	5	3	1	5	7	3	8	2	0
42. Point to middle checker.	16	4	10	5	3	0	6	1	3	12	7	1

Table 5 (cont'd)

Items	++			+-			-+			--		
	U	C	H	U	C	H	U	C	H	U	C	H
43. Point to first checker.	13	5	4	5	2	2	8	3	4	13	5	4
44. Point to last checker.	5	2	9	4	5	0	18	2	3	12	6	2
45. Point to second checker.	1	0	2	3	4	1	10	2	7	25	9	4
46. 2 & 8 - which more?	25	10	11	7	2	0	6	3	3	1	0	0
47. 4 & 6 - which less?	8	6	6	10	5	1	9	2	4	12	2	3
48. 5 & 5 - which more?	0	0	1	0	0	1	5	2	5	34	13	7
49. Which most like wheel?	34	13	13	1	0	0	4	2	1	0	0	0
50. Which most like stick?	29	12	12	2	1	0	7	2	0	1	0	2
51. Which most like tent?	17	3	7	11	5	3	7	3	2	4	4	2
52. Copy (line).	35	14	14	1	0	0	2	1	0	1	0	0
53. Copy (circle).	36	15	10	0	0	2	2	0	2	1	0	0
54. Copy (square).	15	5	6	2	0	0	9	5	2	13	5	6
55. Copy (triangle).	9	2	2	3	2	0	8	3	4	19	8	8
56. What color (black crayon)?	24	8	11	1	0	0	6	3	3	8	4	0
57. What color (red crayon)?	25	7	13	4	1	0	5	2	0	5	5	1
58. Same color as night.	15	4	8	3	2	1	14	5	4	7	4	1
59. Color (circle	13	7	5	7	1	1	6	6	2	13	1	6
60. {yellow	23	6	11	4	0	0	5	4	2	7	5	1
61. Color {square	7	1	4	8	2	2	10	5	4	14	7	4
62. {purple	9	2	8	6	0	3	15	4	1	9	9	2
63. Color {triangle	9	6	4	5	0	1	12	5	5	13	4	4
64. {orange	25	6	10	3	1	1	4	5	2	7	3	1

Declines in performance were much more numerous in the Ulysses group. There were 64 instances of such retrogression in Ulysses as compared with 15 in Clark and 17 in Hillcrest. The declines in performance at Ulysses occurred mostly in some simple directions such as "show me your heel," "wiggle," and particularly in items calling for a pantomime response.

Animal House (WPPSI)

Differences among the groups on pretest scores were relatively small (see Figure 4). The distribution of pretest scores was virtually identical for the Ulysses and Clark groups, with the Hillcrest scores just barely higher. The mean scaled pretest score of 9.5 for both the Ulysses and Clark groups was slightly below the mean of 10 given by the test norms. The mean scaled pretest score of the Hillcrest group was just one-third of a standard deviation above the mean of the standardization group.

Changes in performance from pretest to posttest were relatively small (see Figure 5) and could be accounted for entirely in terms of the increase in age among the children. Increments in scaled scores, i.e., scores scaled as a function of age, were close to zero in all three groups; none of the three school groups showed changes in performance that differed from what might have been obtained in groups that were tested at two points in time without an intervening period of preschool experience.

Matrix Test

Pretest scores on the Matrix Test varied sharply as a function of item difficulty. Almost all the children were able to respond correctly to the first two Perceptual Matching items -- a geometric figure (colored circle) and a pocket watch, but the Clark group was less effective in dealing with the third Perceptual Matching item, which presented a more complex depiction of a cow. All three groups met with widespread failure upon dealing with the fourth Perceptual Matching problem (see Table 6). In this item, the distinctive characteristic of the figure was the relation of its parts to each other, i.e., three pyramids were presented in descending order of size. The lower total scores obtained by the Clark group on the Perceptual Matching items (see Figure 6) are almost entirely attributable to the greater difficulty they experienced with the item presenting the cow (item 3).

Variation in performance on the Class Membership items was produced more by the character of the item than by the school group (see Table 6). However, the Clark group performed less well on most of these items. The Ulysses and Hillcrest groups outperformed the Clark group by a large margin on the first Class Membership item, one in which the common element was the figure's form -- i.e., its triangular shape. Performance in all three groups declined sharply on the following item (item 6), in which the common element was the color of the diversely shaped figures in the matrix.

Several Class Membership items produced successful pretest responses in all the groups. These were items in which the common elements were fruits (item 7), birds (item 8), vehicles (item 10), and houses (item 14). These commonalities were salient defining characteristics of the elements of the matrix. Slightly less inclusive categories -- dogs (rather than four-legged

Figure 4

Distribution of WPPSI Animal House Pretest Raw and Scaled Scores

Raw Score

<u>RAW SCORE</u>	<u>Ulysses (N = 38)</u>	<u>Clark (N = 15)</u>	<u>Hillcrest (N = 14)</u>
50-54	XXX	X	
45-49	XX		X
40-44	X	XX	XX
35-39	XX		X
30-34	XXXX	X	XXXXX
25-29	XXXXX	XX	X
20-24	XXXXXXXX	XX	X
15-19	XXXX	X	XX
10-14	XXXXXX	XXXXX	
5-9	XXX	X	X
0-4	X		
	$\bar{X} = 24.9$	$\bar{X} = 22.5$	$\bar{X} = 29.9$

Scaled Score

<u>SCALED SCORE</u>	<u>Ulysses (N = 38)</u>	<u>Clark (N = 15)</u>	<u>Hillcrest (N = 14)</u>
17	X		
16			
15	XX	X	X
14			XX
13	XX	X	XX
12	XXX	XX	XX
11	XXXXX		XX
10	XXX	XXX	XXX
9	XXXXXXXXX	XXX	X
8	XXXXXXXX	X	
7	X	XXX	
6	XXXX		X
5		X	
4	XX		
3			
2			
1			
0			
	$\bar{X} = 9.5$	$\bar{X} = 9.5$	$\bar{X} = 11.4$

Figure 5

Distribution of Change Scores of WPPSI Animal House Subtest Scores

Raw score

CHANGE SCORE	Ulysses (N = 38)	Clark (N = 15)	Hillcrest (N = 14)
+ 34-36	X		
+ 31-33	X		
+ 28-30		X	
+ 25-27	X		
+ 22-24	X	X	
+ 19-21	XX		X
+ 16-18	XXX	XX	X
+ 13-15	XXXXXX	X	
+ 10-12	X	XX	XX
+ 7-9	XXXX	X	XXX
+ 4-6	XXXXXXXX	XX	XXX
+ 1-3	XX	X	
0			X
- 1-3	XXXX	X	X
- 4-6	XX	X	X
- 7-9		X	X
- 10-12		X	
- 13-15	X		
- 16-18	X		
	$\bar{X} = +8.76$	$\bar{X} = +7.60$	$\bar{X} = +5.86$

Scaled score

CHANGE SCORE	Ulysses (N = 38)	Clark (N = 15)	Hillcrest (N = 14)
+7	X		
+6			
+5	X		
+4	XX	XX	X
+3	XX		
+2	XXXXXX	XX	X
+1	XXXXXX	XXX	XXX
0	XXXXXX	XXX	XXX
-1	XXXXXXX	X	XX
-2	XX		X
-3	XXXX	XXX	XXX
-4	X	X	
	$\bar{X} = +.45$	$\bar{X} = +.07$	$\bar{X} = -.29$

Table 6

Percent of Ss Answering Each Matrix Item Correctly on Pretest

	<u>Item No.</u>	<u>Ulysses</u> (N = 19)	<u>Clark</u> (N = 10)	<u>Hillcrest</u> (N = 12)
Perceptual Matching	1	94	100	83
	2	100	90	84
	3	85	50	83
	4	32	30	41
Class Membership	5	84	50	92
	6	58	30	41
	7	79	70	67
	8	89	60	83
	9	21	20	25
	10	89	80	75
	11	48	60	41
	12	21	20	17
	13	69	40	75
	14	89	60	83
	15	53	40	58
	16	79	60	8
	17	37	30	41
	18	48	20	75
One-Way Classification	19	42	40	50
	20	32	20	25
	21	53	80	41
	22	42	90	59
	23	26	20	8
Two-Way Classification	24	11	10	17
	25	5	0	16

Figure 6

Distribution of Pretest Scores on Four Clusters of Matrix Test Items

	<u>Score</u>	<u>Ulysses</u> (N = 19)	<u>Clark</u> (N = 10)	<u>Hillcrest</u> (N = 12)
Perceptual Matching	4	XXXXX	XX	XXX
	3	XXXXXXXXXXXX	XXX	XXXXXXXX
	2	XXX	XXXXX	
	1			XX
		X = 3.11	X = 2.70	X = 2.92
Class Membership	14-13			
	12	XX		X
	11	X	X	X
	10	XXX		XX
	9	XXXXXX	XX	XXX
	8	XX		X
	7	XX	XX	X
	6	X		X
	5	X	XXX	
	4	X		
	3		X	
	2		X	X
	1			X
		X = 8.63	X = 6.30	X = 7.83
One-Way Classification	5	X		
	4	XXX		XX
	3	X	XXXXXXXX	X
	2	XXXXX	X	XXXX
	1	XXXXXXXX	XX	XX
	0	XX		XXX
		X = 1.95	X = 2.50	X = 1.75
Two-Way Classification	2			
	1	XXX	X	XXXX
	0	XXXXXXXXXXXXXXXX	XXXXXXXXXX	XXXXXXXXXX
		X = .16	X = .10	X = .33

animals) and boys or girls (rather than people) -- produced lesser degrees of success. The most striking difference in performance found among groups on any single item occurred in item 16, in which a picture of a fish had to be recognized as an animal belonging to a matrix containing a horse, a cat, and a dog. Both the Ulysses and Clark Head Start Center groups responded successfully to this item, whereas almost all the Hillcrest children failed it. The more abstract Class Membership items, i.e., those calling for the recognition of large things or things which can be found in the street, were generally more difficult for all the groups but especially for the Clark children.

Pretest performance on the One-Way Classification items was generally poor, and mean scores on these items (see Figure 6) were quite similar among the three groups. The most surprising element of the findings involving One-Way Classification problems was the sudden rise in performance of the Clark group on two One-Way Classification problems involving clearly distinguishable stimuli (representations of children and animals -- items 21 and 22) after rather mediocre performance on two similar preceding problems which presented abstract geometric figures (items 19 and 20). This almost perfect performance in Clark on the representational items was not matched by the other two groups, though their overall performance on the One-Way Classification problems was on a par with that of the Clark group. As has been found in previous studies involving the Matrix Test, the shift from One-Way Classification problems involving vertical arrays to those involving horizontal arrays (item 23) further reduced the incidence of success. As has also been found in the past, there were only rare instances of success on the Two-Way Classification problems.

It is of interest that changes in Matrix Test performance from pre- to posttest (see Figure 7) were relatively small in all three groups. However, among the Perceptual Matching items, there were large gains made in the ability to deal with the most difficult item -- the row of pyramids in descending order -- in both the Clark and Hillcrest groups but not in the Ulysses group. For the most part, gains in performance on individual items did not exceed declines by a substantial margin. Among the Class Membership items it was only the Clark group, which had lagged in pretest performance, that showed consistent though modest gains. Among the One-Way Classification problems, the Hillcrest group showed the greatest improvement.

Kaleidoscope

Differences in performance on the Kaleidoscope task during pretesting among the three groups were negligible (see Figure 8). The pattern of change scores based on pre-post differences was also relatively similar among the three school groups (see Figure 9). This apparent measure of curiosity behavior was insensitive to differences among the groups.

Figure 7

Distribution of Guess Scores on Four Clusters of Matrix Test Items

	<u>Guess Score</u>	<u>Wynnes (N = 12)</u>	<u>Clark (N = 10)</u>	<u>Hillcrest (N = 12)</u>
Perceptual Matching	+3			X
	+2		X	
	+1	XX	XXX	XXXXX
	0	XXXXXXXXXX	XX	X
	-1	XXXX	XX	XXX
	-2	XX	XX	
	-3			XX
		X = -.32	X = -.10	X = -.10

Class Relationship	+5	X	X	
	+4		X	
	+3		XX	XX
	+2	X	X	
	+1	XXXX	XX	XX
	0	X	XXX	XX
	-1	XXXX		XX
	-2	XXX		XX
	-3	XX		XX
	-4	X		
	-5	X		
		X = -.68	X = +1.50	X = -.33

One-Way Classification	+4			X
	+3	X	X	X
	+2	XXX		XX
	+1	XXX	X	XXX
	0	XXXXX	XXXXX	XX
	-1	XXXX	XX	XXXX
	-2	XXX		
		X = +.11	X = -.20	X = +.58

Two-Way Classification	+2			X
	+1	X		XXX
	0	XXXXXXXXXXXX	XXXXXXXXX	XXXXX
	-1	XXX	X	XXX
		X = -.11	X = -.30	X = +.17

Figure 8

Distribution of Pretest Kaleidoscope Scores*

Scores based on first ten items (maximum score = 10)

<u>SCORE</u>	<u>Ulysses</u> <u>(N = 20)</u>	<u>Clark</u> <u>(N = 10)</u>	<u>Hillcrest</u> <u>(N = 12)</u>
8-10			
7	X	X	X
6	X	XX	
5	XXXX	XX	XXX
4	XXXX		XX
3	XXXX	X	X
2	X	XX	XX
1	XXXX		X
0			X
	$\bar{X} = 3.6$	$\bar{X} = 4.3$	$\bar{X} = 3.6$

Scores based on all 14 items (maximum score = 18)

<u>SCORE</u>	<u>Ulysses</u> <u>(N = 20)</u>	<u>Clark</u> <u>(N = 10)</u>	<u>Hillcrest</u> <u>(N = 12)</u>
12-14			
11	X	X	X
10		X	
9			X
8	XX	X	
7	XX	X	X
6			XX
5	XXXXX	XX	XX
4	XXX	X	XX
3	XXX	X	X
2	X	XX	
1	XXX		X
0			X
	$\bar{X} = 4.6$	$\bar{X} = 5.7$	$\bar{X} = 5.1$

*The first 10 items in this measure have a maximum score of 1 each; items 11 to 14 have a maximum score of 2 each.

Figure 9

Distribution of Pre-Post Kaleidoscope Change Scores

Changes in score based on first ten items

CHANGE SCORE	Ulysses (N = 20)	Clark (N = 10)	Hillcrest (N = 12)
+5		X	
+4	X	X	
+3		X	X
+2	XXXXXXX	XXX	XXX
+1	XXX		XXXXXXX
0	XXXX		X
-1	XX	X	
-2	X	X	
-3	XX	X	
-4			
-5		X	
	$\bar{X} = .55$	$\bar{X} = .90$	$\bar{X} = 1.30$

Changes in score based on all 14 items

CHANGE SCORE	Ulysses (N = 20)	Clark (N = 10)	Hillcrest (N = 12)
+10	X		
+9			
+8		XX	X
+7			
+6	XXXXX	X	X
+5		XX	XXX
+4	XX	X	X
+3	XX		XX
+2	XXX		
+1	XX	X	XX
0	X		X
-1	XX		X
-2	X	X	
-3	X		
-4			
-5			
-6			
-7		XX	
	$\bar{X} = 2.7$	$\bar{X} = 2.1$	$\bar{X} = 3.3$

Figure Drawings

The results of the analysis of pretest figure drawings are presented in Table 7. While most of the drawings obtained during the pretest period were self-depictions, there were many exceptions to this pattern. Virtually all the Hillcrest drawings were self-depictions but in the Ulysses group only a little more than half the drawings were of the self. Perhaps more to the point was the degree to which what was drawn was recognizable. All the Hillcrest drawings were recognizable, in contrast to the 60 and 70 percent of the Ulysses and Clark drawings, respectively, which fell in this category. It is of interest that while the Hillcrest children more often made more recognizable drawings, they did not more often draw elaborated central figures. Relatively few drawings contained central figures that were elaborated; the fewest came from the Hillcrest group.

Application of the Harris-Goodenough scoring system to the drawings pretest revealed substantial differences between the Head Start center groups and the middle-class nursery school group in mean raw score (see Table 8). These differences became larger when the IQ was computed, thereby crudely correcting for the difference in age among the groups. When their drawings were scored according to the Harris-Goodenough method, relatively few children from the Ulysses and Clark samples achieved a level exceeded by most of the Hillcrest children.

A number of different qualitative analyses of the drawings often revealed similar differences among the groups but of lesser magnitude. The Hillcrest drawings tended to be rated more complete -- even when the unrecognizable drawings were discounted (see Table 9) -- and their placement and size were rated as better. There was a smaller difference, favoring the Hillcrest group, with regard to the quality and function of the lines that were drawn and also with regard to the difficult judgment of the degree to which the realization of the intent in the drawing was achieved and the degree of organization and unity in the drawing. Manifestations of affect were most often found in the Hillcrest drawings and they were rated as having slightly more vigor and expressiveness.

The major shift in performance of the Head Start children from pretest to posttest was that virtually all their drawings were recognizable in the posttest. There were relatively few changes in any of the three groups toward an elaborated central figure.

Slightly greater gains in Harris-Goodenough scores were obtained in the Clark and Hillcrest groups than in the Ulysses group (see Table 10). On the whole, ratings of the drawings showed almost as many instances of decline in performance as rises; the mean change scores were usually positive but very low. Changes in ratings of the degree of completeness of the human figure between pretest and posttest performance were more variable for the Head Start children than for the Hillcrest group but overall mean differences among the groups were negligible. Changes in the ratings of placement and size, quality and function of line, and realization of intent for the most part clustered around the zero change score in all three groups.

Table 7

Qualitative Features of Pretest Figure Drawings

	<u>Ulysses</u> <u>(N = 22)</u>	<u>Clark</u> <u>(N = 10)</u>	<u>Hillcrest</u> <u>(N = 12)</u>
<u>Content of Drawing</u>			
self-depiction	12	7	11
person other than self	0	0	1
other or unknown	10	3	0
<u>Recognizability</u>			
depiction recognizable	13	7	12
depiction unrecognizable	9	3	0
<u>Number of Figures</u>			
1 figure only	14	8	11
2 figures only	1	1	1
more than 2 figures	1	1	0
cannot judge	6	0	0
<u>Elaboration of Central Figure</u>			
central figure elaborated	5	2	1
central figure not elaborated	9	6	11
cannot judge	8	2	0

Table 8

Harris-Goodenough Scores of Pretest Figure Drawings:
Mean and Standard Deviation

	<u>Ulysses</u> <u>(N = 22)</u>		<u>Clark</u> <u>(N = 10)</u>		<u>Hillcrest</u> <u>(N = 12)</u>	
	<u>$\bar{X}$</u>	<u>s</u>	<u>$\bar{X}$</u>	<u>s</u>	<u>$\bar{X}$</u>	<u>s</u>
Harris-Goodenough Adapted Score	67.91	13.66	67.50	12.41	81.92	12.41
Harris-Goodenough Raw Score	4.41	4.37	3.90	4.25	8.17	4.09
Harris-Goodenough Standard Score	3.86	3.47	3.20	2.89	7.00	2.98

Table 9

Ratings of Pretest Figure Drawings:
Mean and Standard Deviation*

	Ulysses (N = 22)		Clark (N = 10)		Hillcrest (N = 12)	
	$\bar{X}$	s	$\bar{X}$	s	$\bar{X}$	s
1. Completeness of human figure	3.68	3.15	3.70	3.12	7.17	2.24
2. Placement and size	6.59	1.84	6.80	2.04	7.67	2.34
3. Quality and function of line	6.18	1.78	6.20	1.13	7.17	1.58
4. Realization of intent, organization, unity	6.18	2.40	6.20	1.54	7.67	1.92
5. Affect: presence or absence	3.82	.90	3.70	.48	4.50	.90
6. Expressiveness, vigor	2.68	.71	2.40	.70	3.00	.85

*These figures are based on the sum of two independent sets of ratings. The possible range for the first scale is 1-10, for the next three scales 2-10, and for the last two 2-6 and 2-4 respectively.

Table 10

Mean Pre-Post Changes in Figure Drawing Scores and Ratings

	<u>Ulysses</u> <u>(N = 22)</u>	<u>Clark</u> <u>(N = 10)</u>	<u>Hillcrest</u> <u>(N = 12)</u>
Harris-Goodenough Adapted Score	+ 1.59	+ 2.60	+ 1.58
Harris-Goodenough Raw Score	+ 1.68	+ 2.30	+ 3.00
Harris-Goodenough Standard Score	+ 1.55	+ 4.50	+ 3.42
Completeness of human figure	+ 1.27	+ 1.80	+ .25
Placement and size	+ .86	+ .20	+ .08
Quality and function of line	- .27	+ .50	- .08
Realization of intent, organization, unity	+ .32	+ .20	+ .17
Affect: presence or absence	+ .41	.00	- .17
Expressiveness, vigor	.00	- .20	+ .17

Observation of Individual Children

At the conclusion of each cycle of two fifteen-minute observations of a single child, the observer was asked to make a series of judgments with respect to the following categories of behavior: purposefulness of child's behavior, mode of mastery re materials, mastery of school environment, frustration tolerance, impulse control, mode of social relationships with children, curiosity, and strength of self (see Appendix for a more detailed presentation). More often than not the observers felt that their 30-minute observation was too brief to provide conclusive information regarding the behavior in question. As a result, they often lacked faith in the validity of their judgments and in a number of instances could make no judgment because they lacked a firm basis for doing so. Nevertheless, the results of these judgments are presented here, despite the unreliability resulting from lack of evidence or incidence, or from differences among observers' judgments (see Table 11).

Purposefulness of child's behavior. In all three groups, most of the child's behavior was not regarded as random or aimless. The children from the Hillcrest school were seen as more socially motivated.

Mode of mastery re materials. While most of the children were seen as functioning comfortably and independently, this was most true of the Hillcrest group. The Hillcrest group was seen to be the most deeply involved in process, whereas the Clark children were judged to be somewhat more end-product oriented. The Hillcrest children were regarded as least distractable.

Mastery of school environment. Most of the children were viewed as knowing where to find things and put things back and how to get things that they could not find.

Frustration tolerance. There were more instances in the two Head Start groups than in Hillcrest of the child's frustration stemming from limits imposed, e.g., requirements to wait, to share, or to sit still. The response to frustration more often involved throwing and hitting in the Clark school and there were fewest instances of withdrawal or passive acceptance as a response to frustration in the Hillcrest school. The Clark children were also seen more often as likely to ask for help when in difficulty.

Impulse control. There were more cases of excessively controlled children in the two Head Start programs. Most children were seen as adequately controlled and their behavior well-integrated; this pattern was especially prevalent in the Hillcrest group.

Mode of social relationships with children. The most frequent mode of approach to other children that was observed was that of conversing with them. The Clark children, more often than the other groups, resorted to teasing and threatening as an overture to other children. In general, more physical methods of approaching other children were used in the Head Start groups. Constructive, friendly approaches to the children were most often found among the Hillcrest children.

U = Ulysses (N = 12)
 C = Clark (N = 12)
 H = Hillcrest (N = 12)
 + = Yes
 - = No
 0 = No evidence

- 20a -

Table 11

Summary of Observer's Judgments on Aspects of Children's Behavior
 Based on Two Observation Cycles

	U			C			H		
	+	-	0	+	-	0	+	-	0
A. PURPOSEFULNESS OF CHILD'S BEHAVIOR									
Child's behavior is random or aimless most of the time	4	15	4	4	13	6	1	23	-
Child is involved in a situation primarily because of a desire to be with a particular friend or group, i.e., he is socially motivated	9	9	5	10	11	2	14	4	6
B. MODE OF MASTERY RE MATERIALS									
Child functions comfortably and independently	10	5	7	13	9	1	20	3	1
Child gets very involved in process and is not concerned much about end product	12	4	7	8	4	11	15	4	5
Child is end-product oriented	2	13	8	4	7	12	-	14	10
Child seems quite competent but his attention span is very limited and he is very easily distracted	4	6	13	4	5	14	-	16	8
C. MASTERY OF SCHOOL ENVIRONMENT									
Child knows where to find what he needs	16	-	7	17	-	6	21	-	3
Child knows where to put things back	15	-	8	16	-	7	14	-	10
Child knows how to get what he can't find (e.g., asks teacher or another child)	8	-	15	14	-	9	13	-	11
D. FRUSTRATION TOLERANCE									
<u>Possible Sources of Frustration</u>									
Limits imposed by adult (e.g., child asked to stop, to wait, etc.)	5	8	10	5	1	17	3	9	12
Enforced sharing (often experienced as deprivation)	3	2	18	6	-	17	-	6	18
Enforced stillness (e.g., sitting or listening for too long)	1	3	19	4	-	19	-	5	19

Table 11. (cont'd)

	U			C			H		
	+	-	0	+	-	0	+	-	0
D. FRUSTRATION TOLERANCE (cont'd)									
Possible Responses to Frustration									
Throwing, hitting, kicking, etc.	1	11	11	6	3	14	-	8	16
Withdrawal	7	8	8	5	5	13	-	8	16
Passive acceptance (is situation resolved?)	9	4	9	6	5	12	3	7	14
Asking for help when in difficulty	4	7	12	11	2	10	2	5	17
E. IMPULSE CONTROL									
Excessive control (child who is seemingly "overcontrolled" -- who may or may not be withdrawn or docile)	8	7	8	6	13	4	1	22	1
Integrated behavior (adequate control)	12	4	6	15	4	4	23	-	-
F. MODE OF SOCIAL RELATIONSHIPS WITH CHILDREN									
Child's Approach to Other Children									
Verbal-constructive: suggestions, comforting, help, etc.	9	8	6	5	2	16	13	3	8
Verbal-aggressive: teasing, threatening, etc.	4	10	8	9	4	10	3	12	9
Verbal-neutral: talking, neutral conversation	14	5	4	18	1	3	21	2	1
Physically affectionate: with or without physical contact	9	8	6	11	3	9	12	3	9
Physically aggressive: hitting, destroying, interfering with play or products, removing things	6	9	7	9	4	10	-	15	9
Neutral physical participation: just joining the activity, offering a thing	14	3	5	12	2	9	14	5	5
Child's Response to Other Children's Approaches									
Verbal-constructive: suggestions, comforting, help, etc.	6	7	10	6	2	15	13	3	8
Verbal-aggressive: teasing, threatening, etc.	2	11	10	8	1	14	6	12	6
Verbal-neutral: talking, neutral conversation	16	3	4	19	-	4	23	1	-

Table 11 (cont'd)

	U			C			H		
	+	-	0	+	-	0	+	-	0
Child's Response to Other Children's Approaches (cont'd)									
Physically affectionate: with or without physical contact	10	6	7	11	3	9	12	2	10
Physically aggressive: hitting, destroying, interfering with play or products, removing things	6	8	9	8	2	13	-	17	7
Neutral physical participation: just joining the activity, offering a thing	13	3	7	13	-	10	13	5	6
G. CURIOSITY									
Does the child notice and react to change and new events in the familiar setting (e.g., does he comment on different pictures on the wall, the absence of a particular child, a new toy on the block shelf, that there is a different kind of juice, etc.)?	6	1	16	6	1	16	8	-	16
Does the child respond with interest to and participate in new experiences offered to the child or group by the teacher (e.g., to new materials offered in science corner, to events and sights of a walk, to new books, etc.)?	8	-	15	5	1	17	17	-	7
Does the child ask questions to find out about things and situations from others?	3	9	11	7	3	13	11	-	13
H. STRENGTH OF SELF									
Is he aware of what he wants?	12	1	10	15	1	7	19	-	5
Does he assert his own rights?	10	3	4	13	3	7	15	3	6
Is he competent?	10	5	8	15	-	8	21	1	2

Curiosity. Often the observers believed that they had insufficient evidence to judge the children's curiosity. However, the Hillcrest children were more often regarded as responding with interest to experiences introduced by the teacher. The Clark and Hillcrest children were more likely to ask questions.

Strength of self. There was more evidence of self-assertiveness and awareness of one's needs in the Clark and Hillcrest groups. More of the Hillcrest children were judged to be competent.

In another approach to the analysis of the observation data, the four fifteen-minute observation records gathered for each child were pooled and then each child was rated along a set of dimensions (see Table 12).⁸ Within the overall pattern of substantial similarity among the mean ratings for the four groups, some differences are discernible. The Hillcrest group was judged to be substantially more concentrated, skilled and cooperative, and to show better control than the two Head Start groups. The largest differences between the two Head Start groups involved those which showed the Clark group to be more watchful, sociable, cooperative and more animated than the Ulysses group.

Table 12

Mean Ratings Based on Observation Records of Individual Children
(Scale Range: 1-10)

	<u>Ulysses</u> <u>(N = 12)</u>	<u>Clark</u> <u>(N = 12)</u>	<u>Hillcrest</u> <u>(N = 12)</u>
concentrated-distractable	6.4	6.0	7.4
purposeful-aimless	5.5	5.1	5.4
watchful-unnoticing	5.1	6.4	6.0
verbal-silent	5.2	5.8	6.0
skilled-awkward	5.8	6.4	7.3
meticulous-casual or destructive	6.0	5.8	6.0
proud of work-indifferent	6.1	5.9	6.5
sociable-isolated	5.4	6.6	6.5
cooperative-solitary or fractious	5.3	6.3	6.8
dominant-submissive	5.2	6.0	5.5
independent-dependent	5.1	5.0	4.8
explosive-overcontrolled	4.7	4.8	4.9
animated-unanimated	4.7	7.0	5.7
tense, agitated-phlegmatic	5.6	6.1	5.6

8. These ratings were made by Lois Chaffee.

Table 14

Mean Statements to Teacher and to Peers

	<u>Ulysses</u> <u>(N = 10)</u>	<u>Clark</u> <u>(N = 10)</u>	<u>Hillcrest</u> <u>(N = 5)</u>
<u>Pretest Statements</u>			
to Teacher	6.2	10.2	13.0
to Peers	48.9	38.3	38.6
<u>Posttest Statements</u>			
to Teacher	8.2	13.0	14.4
to Peers	73.0	55.4	46.4

The Clark children were often than those from Ulysses asked for things they needed (designated as Desire Implementation in the category system, presented in Table 15).

Table 15

Mean Frequency of Desire Implementation Statements

	<u>Ulysses</u> <u>(N = 10)</u>	<u>Clark</u> <u>(N = 10)</u>	<u>Hillcrest</u> <u>(N = 5)</u>
Pretest	6.5	8.6	16.4
Posttest	9.2	8.3	6.8

The Hillcrest children were scored as asking more than twice as many requests as the Ulysses group at the outset, with this number declining sharply by posttest. It is interesting that in all three groups the frequency of this category decreased from pre- to posttesting.

The Clark group increased markedly in socially-oriented talk from pre- to posttesting (see Table 16), more than doubling their mean frequency of questions (6.4 to 13.4) and answers to peers (1.7 to 4.8) from pre- to posttesting. The Ulysses means remained relatively unchanged from pre- to posttesting. The values were lowest among the Hillcrest children.

Table 16

Mean Frequency of Questions and Answers to Peers

	<u>Ulysses</u> <u>(N = 10)</u>	<u>Clark</u> <u>(N = 10)</u>	<u>Hillcrest</u> <u>(N = 5)</u>
<u>Pretest</u>			
Questions	9.3	6.4	3.4
Answers	2.8	1.7	1.4
<u>Posttest</u>			
Questions	9.6	13.4	8.4
Answers	2.2	4.8	.8

12. Answers to teachers were considered teacher-directed and were not scored.

A cluster of "academic talk" categories were studied in which the frequency of Clark responses was usually twice that of the Ulysses group (see Table 17). This cluster includes statements of pride in acquiring knowledge, statements denigrating the competence or power of peers, statements playing the teacher role in the domain of competence, statements reporting on self-products or productions, and statements implementing learning.

Table 17

Mean Frequency of Academic Talk

	<u>Ulysses</u> <u>(N = 10)</u>	<u>Clark</u> <u>(N = 10)</u>	<u>Hillcrest</u> <u>(N = 5)</u>
<u>Pretest</u>			
Pride in knowledge	.0	.3	.0
Denigrate peer competence	.2	1.1	.2
Assume teacher role-competence	.5	1.1	.2
Report own products	.7	.5	3.2
Implement learning	<u>1.4</u>	<u>1.6</u>	<u>1.8</u>
TOTAL	2.8	4.6	5.4
<u>Posttest</u>			
Pride in knowledge	.0	1.0	.0
Denigrate peer competence	1.0	2.0	1.2
Assume teacher role-competence	2.6	5.4	.0
Report own products	1.6	2.6	2.8
Implement learning	<u>2.4</u>	<u>3.2</u>	<u>1.2</u>
TOTAL	7.6	14.2	5.2

At pretesting, the frequency of these categories was highest in the Hillcrest group, but whereas these frequencies did not change at posttesting in Hillcrest, they tripled in the Head Start groups.

Tables 18 and 19 are based in part on the sub-category scores from which the academic cluster derives. It can be seen in Table 18 that the Ulysses group exceeded the Clark group in the total number of Positive Ego-Enhancing statements. The Pride in Power scores cover Pride in the Whole Self ("I'm big"); Pride in Possessions ("I have a new shirt"); and Pride in Competence and General Achievement ("I can jump"; "Look at my big building"); as well as Pride in Knowledge ("That's a five" */proudly/*). The Pride in Competence and General Achievement, the most frequently occurring source of pride, was somewhat higher for the Ulysses group in both pre- and posttesting. The Hillcrest group scored highest of all and resembled the Ulysses pattern quite closely. The Pride in Possessions category was more prominent in the Hillcrest group.

Table 18

Mean Frequency of Positive Ego-Enhancing Statements

	<u>Ulysses</u> <u>(N = 10)</u>	<u>Clark</u> <u>(N = 10)</u>	<u>Hillcrest</u> <u>(N = 5)</u>
<u>Pretest</u>			
Pride in Power			
whole self	.3	.2	.2
possessions	.9	.6	2.0
competence and general			
achievement	4.6	3.5	4.8
knowledge	.0	.3	.0
Pride in Goodness	1.0	.4	.2
TOTAL	6.8	5.0	7.2
<u>Posttest</u>			
Pride in Power			
whole self	.0	.4	.0
possessions	2.8	3.2	4.0
competence and general			
achievement	7.8	5.0	8.0
knowledge	.0	1.0	.0
Pride in Goodness	1.6	.6	.8
TOTAL	12.2	10.2	12.8

The Self-Report category findings presented in Table 19 include Report Possessions and Attributes ("My hair is curly"), Reporting Doing ("I'm playing with the blocks now") and Reporting on Products and Productions ("This is the Shell gas station"). At posttesting, the Clark focus was on reporting self-products or productions, while the Ulysses group maintained its original emphasis on reporting what the self was doing. Other self-report scores were more evenly divided between reporting attributes and possessions and reporting self-products and productions.

Table 19

Mean Frequency of Statements Reporting on the Self

	<u>Ulysses</u> <u>(N = 10)</u>	<u>Clark</u> <u>(N = 10)</u>	<u>Hillcrest</u> <u>(N = 5)</u>
<u>Pretest</u>			
Possessions and Attributes	.4	.2	1.2
Doing	1.4	1.2	1.6
Products	.7	.5	1.2
TOTAL	2.5	1.9	4.0
<u>Posttest</u>			
Possessions and Attributes	1.2	.6	2.8
Doing	2.4	1.0	4.8
Products	1.6	2.6	2.8
TOTAL	5.2	4.2	10.4

It is interesting that, as with Pride statements, the Ulysses group received a somewhat higher overall Self-Report score, even when reporting on products is included. The most dramatic finding, however, is the much higher frequency of this category in the Hillcrest group.

While the Clark teachers seemed to foster teacher-oriented, question-answer-oriented, academic-oriented talk, the Ulysses teachers tended to emphasize peer-oriented play-type talk. Table 20 shows the results for what is called the Collaborative category. This category includes talk when S is engaged in a collaborative role-differentiated project or discussion, talk in a dramatic role-playing sequence, collaborative chanting games and nurturant, giving statements.

Table 20

Mean Frequency of Collaborative Talk

	<u>Ulysses</u> <u>(N = 10)</u>	<u>Clark</u> <u>(N = 10)</u>	<u>Hillcrest</u> <u>(N = 5)</u>
<u>Pretest</u>			
Projects and discussions	16.4	8.8	7.4
Dramatic play	5.8	3.0	7.4
Chanting	.9	1.7	1.8
Giving	3.1	1.0	1.4
TOTAL	26.2	14.5	18.0
<u>Posttest</u>			
Projects and discussions	15.0	12.4	17.2
Dramatic play	9.4	4.0	2.0
Chanting	6.8	4.4	6.4
Giving	1.6	.4	2.4
TOTAL	32.8	21.2	28.0

It can be seen that the Ulysses group exceeded the Clark group on both pre- and posttesting in this collaborative category. Except for the chanting category on pretesting, all four kinds of statements show Ulysses exceeding Clark on pre- and posttesting. The Hillcrest group occupies a point between the two Head Start groups on these scores. It may also be noted that this is a very high frequency category, accounting for 25 to 30 percent of all speech in the total group. It is also interesting to note the findings on dramatic role-playing talk. This was strictly defined as talk while enacting a role. The Clark group increase of mean statements (from 3.0 to 4.0) can be contrasted with the almost doubling (from 5.8 to 9.4) of the Ulysses means.

Table 21 shows the mean number of times S had to stop a frustrator or defend himself against someone who was stopping him as a frustrator. This Frustration Talk was stated generally ("Stop it," "No"), or specifically, in terms of possession rights ("It's mine," "I had it first"). It can be seen that the two groups did not differ very much in their overall frustration talk, with Clark 1.6 points ahead at pretesting and Ulysses 1.4 points

ahead at posttesting. However, it is clear that the Clark group showed a distinct increase in appeal to possession rights which the Ulysses group did not show. This increase may well be a by-product of the Montessori structure of materials in the Clark program; it may be easier to perceive a distinctive single piece of Montessori equipment one is working with as "mine," than one of many dolls, blocks, or cars, when another child comes to grab it. The frequency of frustration statements was considerably lower in the Hillcrest group at both pre- and posttest.

Table 21

Mean Frequency of Frustration Talk

	<u>Ulysses</u> <u>(N = 10)</u>	<u>Clark</u> <u>(N = 10)</u>	<u>Hillcrest</u> <u>(N = 5)</u>
<u>Pretest</u>			
General	4.4	5.3	2.0
Possession right specified	<u>1.3</u>	<u>2.0</u>	<u>2.6</u>
TOTAL	5.7	7.3	4.6
<u>Posttest</u>			
General	9.4	5.6	2.8
Possession right specified	<u>2.8</u>	<u>6.2</u>	<u>4.0</u>
TOTAL	12.2	11.8	6.8

Tabulation of the category of Angry Talk (see Table 22) indicates that the Clark group exceeded the Ulysses group at pretesting and rose to more than twice as much angry talk at posttesting. Angry Talk includes denigrations and defense against denigrations; exclusion and defense against exclusion; teasing or testing limits (a modulated form of angry expression); and name-calling and cursing. The latter is placed in brackets because it is an appended score, qualifying one of the other scores. It can be seen that the Clark group exceeded the Ulysses group in every kind of angry talk. The Hillcrest group occupied a point between the two Head Start groups at pretesting, mainly because of the larger amount of teasing that took place there, but more closely resembled the rate shown by the Ulysses group at posttesting.

Table 22

Mean Frequency of Angry Talk

	<u>Ulysses</u> <u>(N = 10)</u>	<u>Clark</u> <u>(N = 10)</u>	<u>Hillcrest</u> <u>(N = 5)</u>
<u>Pretest</u>			
Denigration	2.2	3.6	1.0
Exclusion	1.4	1.2	.4
Teasing	.4	1.7	3.4
(name calling-cursing)	<u>(.0)</u>	<u>(.3)</u>	<u>(.8)</u>
TOTAL	4.0	6.5	5.6
<u>Posttest</u>			
Denigration	3.2	7.6	3.6
Exclusion	1.4	2.6	.0
Teasing	2.4	4.8	4.8
(name calling-cursing)	<u>(.4)</u>	<u>(1.0)</u>	<u>(.0)</u>
TOTAL	7.0	15.0	8.4

While the quantity of angry talk at Clark may reflect a negative response to the pressures of the program, there are alternative explanations. The Clark community was more representative of extremes of poverty than the Ulysses community, and this difference may be reflected in the angry talk patterns.

The Hillcrest group was different from both Head Start groups in the use of the "Me Too" category. This category covers a variety of statements where S hears another child's statement and then makes a self-reference related to what the other child has said. The self-reference may involve an identical statement (S₁: "I'm making a monster"; S₂: "I'm making a monster too") or a similar one (S₂: "I'm making a giant") or one that involves a difference (S₁: "My milk comes in containers"; S₂: "Mine comes in bottles") or the self-reference might be tangentially related (S₁: "Look at the fish"; S₂: "My uncle has a fish tank"). Sometimes the words "Me too" are stated. Often the statements end with "too."

Table 23 indicates that Me Too statements occurred much more frequently in the middle-class group, about twice as much at posttesting, than in the Head Start groups. It is interesting that this kind of conversing and connecting with others occurred minimally in the Clark group at pretesting while the Ulysses group showed almost as much as the Hillcrest group at pretesting. However, the Hillcrest group far exceeded the Ulysses group at posttesting, while both Head Start groups tended to equalize.

Table 23

Mean Frequency of Me Too Statements

	<u>Ulysses</u> <u>(N = 10)</u>	<u>Clark</u> <u>(N = 10)</u>	<u>Hillcrest</u> <u>(N = 5)</u>
Pre	3.9	1.6	5.2
Post	5.0	4.0	10.4

Table 24 presents the frequency of statements that were qualified by taking into consideration the viewpoint of the hearer. The Hillcrest group far exceeds either Head Start group in these qualified statements. They include Permission-Please statements which request tentatively ("Can I," "May I," or with "please"); Modulations, statements involving explanations, justifications, rationalizations, or persuasions (sometimes including "because"); Sharing and Postponing, suggested as a solution to a conflict; Collaborative Giving, nurturant, generous statements. The Hillcrest group had almost twice as many of these statements as the Head Start groups. Substantial gains were recorded in this category in Clark but not Ulysses.

Table 24

Mean Frequency of Statements Qualified to Take
Into Consideration the Viewpoint of the Hearer

	Ulysses (N = 10)	Clark (N = 10)	Hillcrest (N = 5)
<u>Pretest</u>			
Permission-please	.9	.8	3.8
Modulation	1.5	1.2	2.4
Sharing and postponing	.6	.7	1.6
Collaborative giving	<u>3.1</u>	<u>1.0</u>	<u>1.4</u>
TOTAL	6.1	3.7	9.2
<u>Posttest</u>			
Permission-please	.8	.6	2.4
Modulation	2.2	4.6	5.6
Sharing and postponing	1.2	1.8	2.4
Collaborative giving	<u>1.6</u>	<u>.4</u>	<u>2.4</u>
TOTAL	5.8	7.4	12.8

Table 25 shows the mean frequency of Orders and Threats in the pursuit of desires and rights. (Frustration Talk is omitted because it was almost always stated in the form of an order.) It can be seen that the Hillcrest group scores were near zero, while both Head Start groups made statements which fell into these categories.

Table 25

Mean Frequency of Orders and Threats¹²

	Ulysses (N = 10)	Clark (N = 10)	Hillcrest (N = 5)
<u>Pretest</u>			
Order	1.6	1.9	.2
Threat	<u>1.1</u>	<u>1.2</u>	<u>.0</u>
TOTAL	2.7	3.1	.2
<u>Posttest</u>			
Order	1.8	1.6	.4
Threat	<u>.0</u>	<u>.6</u>	<u>.0</u>
TOTAL	1.8	2.2	.4

12. In the Desire Implementing and Rights Implementing category, exclusive of Stopping a Frustrator, which was almost always an order.

Discussion

The results of this study are extraordinarily difficult to interpret because of the mass of uncontrolled factors which served to complicate what was from the beginning an imperfect experimental design. The complications themselves, in a sense, become one of the major results of the study, because they suggest that there is an inevitability to the barriers, multifaceted and unforeseeable, which erode efforts to study the influence of contrasting Head Start programs. While the present study was additionally burdened by the fact that it served two masters -- 1) the hypotheses regarding the comparative influence of School Readiness and Child Development approaches to the teaching of Head Start children, and 2) the requirements of the national evaluation study of Head Start in which the study was embedded -- the methodological problems inherent in this area of study are, in themselves, extremely formidable.

Without launching into a comprehensive analysis of methodological problems, it is worth noting some basic facts. The task of instrumentation for the study of the influence of varying programs on the participating children is massive, not only because methods of measurement of the various phenomena -- teacher behavior, classroom atmosphere, group and individual child behavior -- are so underdeveloped in general, but because these methods need to be honed to fit the particular situation. The nature of the atmospheric, teacher, and child variables to be studied can only be established with finality when the actual conditions and subjects of the study are identified. As long as the Head Start program remains relatively unstable, and there are valid reasons for maintaining some of its current instability, it will not be possible to have full and accurate information about the essential ingredients of a program until it is actually underway, and even then, unfortunately, one must be prepared for profound changes associated with the numerous shifts in staffing and the considerable attrition of children. Since many months (if not years) are needed from the moment that full information is available regarding the basic conditions and elements of the study and the completion of suitable instrumentation, a comprehensive pre-post study and of Head Start within the time span of a year is virtually impossible. Further, if a study is to have any generality of findings (a reasonable enough aspiration but one which probably needs to be set aside for the moment), it needs to be conducted in a large number of settings, thereby enormously compounding the problems of instrumentation and effective research administration.

Turning now to the problems which beset this particular study, the results need to be interpreted within a context which considers the fact that suitable exemplars of the School Readiness and Child Development methods could not be found for study, that the classes and children that were contrasted were too few in number, and that their character changed -- in the Head Start groups particularly, teachers and children either left or were shifted to other classes in the course of the year. In addition, delays in the start of testing, together with the long time it required to complete the pretesting, caused the interval between pre- and posttesting to shrink substantially for some children. Finally, the process of negotiating for field settings, and the sheer magnitude and vicissitudes of the data-gathering process itself, seriously interfered with the refinement of instrumentation and the mechanics of data collection.

The first-mentioned of these problems is an important finding in itself. The difficulty in finding suitable prototypes of the School Readiness and Child Development approaches highlights how unexplicit, underarticulated and lacking in self-awareness are the bulk of Head Start teachers with regard to both their objectives and their methods. While recognizing the multiplicity of goals a Head Start teacher is striving to achieve, and the many uncharted paths she is asked to traverse, it is nevertheless noteworthy that so many of the teachers are diffuse and irresolute about their goals and teaching strategies. Clearly, this sphere of teacher functioning needs to be improved if Head Start is to advance to new levels of effectiveness.

Within this unsettling framework, let us turn to the results of the assessment of the children.

The results of the Stanford-Binet pre-post test analysis indicate that the two Head Start groups failed to gain in mean IQ score whereas the Hillcrest group increased its mean IQ by a considerable margin -- 10 points. The middle-class sample began at a mean level one-third of a standard deviation above the Head Start groups and increased this margin at the conclusion of their year of preschool. While performance on some items improved more than on others, there was no clear-cut pattern to the growth in performance of the Hillcrest group. At the same time, a surprising amount of retrogression in performance was noted in all three school groups, but especially in Ulysses.

An inventory of factors affecting Stanford-Binet test performance indicated that, in the judgment of the examiners, relatively few of the children's performance on the Stanford-Binet was adversely affected by the test situation. At pretest, the largest differences in ratings found among the three school groups were as follows: the Hillcrest children were judged to find the tasks least acceptable and to be less responsive to encouragement and the Clark children were rated as being less persistent and showing more impaired activity levels and verbal expression. The Ulysses children were found to be most at ease socially. Whereas test performance of the Hillcrest children was generally found to be less impaired during posttesting than it was during pretesting, and the Ulysses group was judged to be approximately unchanged, the Clark group's performance was found to be markedly more adversely affected during posttesting. It is important to note that at least some of these differences among school groups and between pre- and posttest performance may be attributable to variations among test examiners both within and between groups.

A somewhat similar pattern of performance to that of the Stanford-Binet was found on the Preschool Inventory. Relatively sizeable gains were made by the Hillcrest group, and also by the Clark group. Here, too, there was a greater number of instances of decline in performance in the Ulysses children.

Performance on the Matrix Test failed to differentiate the school groups. At the outset, most of the children showed a capacity to perform simple perceptual matching tasks and could solve those class membership problems

where the common element was a highly salient attribute and the solution could be achieved through an associative response. The effect of the distinctiveness of stimuli in the arrays comprising a one-way classification problem was especially marked in the Clark group. Changes from pre- to posttest performance were relatively slight in all the groups. The greatest changes were in the increased ability of the children to deal with a more intricate perceptual matching task, the Clark group's improvement in performance on class membership problems, and the beginnings of mastery of the one-way classification problems shown by some Hillcrest children.

A measure of curiosity, based on the exploratory behavior evoked by a kaleidoscope placed before the child at the beginning of testing, failed to indicate marked differences among the groups. Differences between before and after testing were also modest.

The figure drawings of the Hillcrest children were found to be superior to those of the Head Start children. They obtained higher Goodenough scores and their drawings tended to be rated as more complete and their placement and size was judged to be better, along with the quality and function of their lines. Their drawings were also rated as more often having realized their intent. The major observable change in the drawings from pre- to posttesting lay in the fact that virtually all the posttest drawings were recognizable. Qualitative ratings of drawings declined almost as often as they rose in the transition from pre- to posttesting.

The results of the observations of individual children indicated that most had made a satisfactory adjustment to preschool in that they had learned to work with the materials, had become familiar with the school environment and, for the most part, behaved purposefully in the course of their school day. The middle-class group was found to function somewhat more independently, seemed less often to be frustrated by imposed limits, was more free from problems of control, more friendly and less aggressive in social contacts, and more competent. Differences between the two Head Start groups were relatively small; the Clark children were seen to be somewhat more aggressive and assertive, and more work-oriented.

The observation of spontaneous language, based on a small number of children, indicated that while the Hillcrest group seemed to speak more at the outset, all three groups had risen to approximately equal levels of volubility by the end of the year. While most of the children's talk was directed toward peers rather than teachers, this was especially true in Ulysses. The largest number of statements directed toward teachers, as opposed to children, was found in the Hillcrest group. There was much more reporting on the self and many fewer frustration statements in Hillcrest. The greatest amount of collaborative talk was found in Ulysses, the least in Clark. There was much more angry talk in Clark and less in Ulysses than in Hillcrest, and more orders and threats issued in the Head Start groups, especially Clark. There were many more references to "me too" and more statements qualified to take into consideration the viewpoint of the hearer in the Hillcrest group.

The results of the assessment of intellectual functioning present a somewhat uneven picture. Substantial differences were found between the Head Start groups and the middle-class sample on the Stanford-Binet, the Preschool Inventory and the Figure Drawing, but not on the Matrix Test (a test which has been found to differentiate the performance of advantaged and disadvantaged children at older age levels), the WPPSI Animal House subtest and the Kaleidoscope (Curiosity) Test. In only two of these tests, the Stanford-Binet and the WPPSI, are the age gradations of the norms fine enough to begin to estimate changes in levels of performance attributable to preschool without reference to a control group. Only on the Stanford-Binet, and only in the case of the privileged (middle-class) sample, was a sizeable change in performance recorded which could be attributed to their preschool experience. There were also gains in performance on the Preschool Inventory, especially in the Hillcrest and Clark groups, but the norms for this test are not sufficiently fine to gauge the extent of change in relative standing. Gains in Figure Drawing performance, largest in the Hillcrest and Clark groups, were relatively small.

The largest gains in Stanford-Binet performance for all groups were in picture vocabulary, comprehension, opposite analogies, knowledge of what common materials are made of, ability to perform commissions, picture completion, paper folding, copying a square, pictorial similarities and differences, and the ability to fit triangles together to form a rectangle. The largest gains in Preschool Inventory performance were in items requiring the ability to follow commissions involving color and relational terms, and knowledge of the function of teachers, dentists and other figures.

The question arises whether the areas of empirically demonstrated change do in fact accurately represent those areas in which the children had grown the most during their preschool experience. Do the measures of change faithfully portray the actual change which transpired? The question has two parts: (1) are the areas of cognitive change theoretically considered to be produced by preschool sufficiently represented among the items of the tests used to measure change, and do they correspond with the content of those test items which show the most change? and (2) are those changes in test performance which have been recorded an indication of an actual change in ability or information level, or rather a change in the availability of skills or knowledge already present? Insofar as the preschool programs have been concerned with the cultivation of communication skills, relatedness to work, attentiveness, curiosity, orderliness and perseverance, these attributes have not been directly assessed by the tests which were employed. As to specific intellectual content, the specific concepts and knowledge learned during the course of the preschool year, it appears likely that these have been only very scantily sampled in the tests that were administered.

There is striking evidence that at least some of the gains in test performance represent changes in the availability of already existing knowledge. The marked increase from pre- to posttesting in the Preschool Inventory in the number of children at the Hillcrest school who could give their first name upon request is surely an example of how children may fail to reveal knowledge or skills in the course of test performance which are quite solidly within their grasp. The extent to which change

scores in test performance reflect changes in the availability of knowledge in the test situation rather than changes in the actual content of knowledge, or even of its availability in other than test situations, is a matter which requires more intensive study than it has heretofore received. Many observers have become concerned with the peculiar unsuitability of the test as an index of disadvantaged children's level of cognitive functioning because of what Labov¹³ has described as the assymmetric relation between examiner and child in the test situation. The numerous instances of retrogressive behavior recorded on many items, especially in the Ulysses group, provides strong evidence that factors other than actual level of knowledge or skill are affecting test performance.

In this regard, it is interesting to note that the rating of factors affecting Stanford-Binet test performance (see p. 16) did not indicate an improvement in test-taking attitudes among the Head Start groups from pre- to posttesting. While it is customary to assume that children are in an advantaged test-taking position at the point of posttesting by virtue of their previous experience with the form and content of the test and their having grown accustomed to relating to other adults during the interim period, the effects of retesting may take another form. The children may feel freer to express their lack of interest in the test, and the fact that the test is a repetition of one previously taken may diminish their attentiveness. Now that they have grown fond of classroom life, they may resent being removed from their class for the purpose of testing. In addition, they may feel impelled to give exactly the same responses they offered during pretesting (voluntarily or involuntarily), thereby further reducing the possibility for showing change. These considerations emphasize the need for re-examining the significance of a measure of change obtained by way of the test-retest method. In this instance, Jane Loevinger's¹⁴ concern for establishing a theory of the second test seems particularly relevant.

The observations of individual children succeeded in providing an overall picture of the main trends of behavior in each of the groups, and it succeeded in revealing some of the major differences between the privileged and underprivileged groups of children. However, the four fifteen-minute observations were far from sufficient for providing a portrait of each individual's characteristic adjustment to preschool. While the category system that was applied to the records helped to focus attention on important aspects of the child's adjustment and thereby facilitated comparisons among groups within a common framework, a much more intensive observation of each child would be necessary to identify those elements of his personality and social development which are most salient in terms of the dynamics

13. Labov, William. "The Logic of Non-standard English," Georgetown Monographs Series on Languages and Linguistics, Monograph No. 22, 1969.

14. Loevinger, Jane. "Objective Tests as Instruments of Psychological Theory," Monograph Supplement of Psychological Reports, 1957, 3, 635-694.

of his particular preschool experience.¹⁵

The study of spontaneous language appears to offer a useful method for describing significant aspects of classroom interaction. Surprisingly, these data appear to suggest that there is very little difference between advantaged and disadvantaged children in the sheer amount of their verbal output. The low ratio of interchange-with-teacher to interchange-with-peers found in the Ulysses Head Start group as compared with the ratio obtained in Clark and Hillcrest is of special interest. In light of the importance of establishing a sense of relatedness to the significant adult, and the acknowledged need for the teacher to actively initiate and give direction to learning experiences, a tendency to schedule extended sessions of non-directed peer group play as the prime form of classroom activity would be a serious misinterpretation of how a child development-oriented preschool class should be taught.

The study of spontaneous language seemed to capture some of the major differences in social interaction between advantaged and disadvantaged children -- the more physical, aggressive style of relating among the disadvantaged, the greater incidence of teasing and insults, of threats and commands, and conversely, the more cooperative and self-enhancing quality of the mode of relating among the privileged children. These observations tend to underline the need for an articulated theory of ego functioning in young children as the appropriate level of analysis for the study of cognitive as well as personality development in preschool children.

15. In a recent pilot study (conducted by Ann Coolidge of the Bank Street College Head Start Evaluation and Research Center) to identify the main areas of growth in a preschool for disadvantaged children, three teachers were asked to describe how each child in their class had changed during the course of the year. Among the changes most often cited were: an increase in assertion among children who were originally withdrawn and retiring; an increase in the ability to follow rules with corresponding decreases in overt aggressive behavior toward strangers and thwarting adults; a reduction of destruction of equipment accompanied by an ability to abide by rules and communicate them to others; a reduction of diffuse anger, alleviation of acute separation anxiety, and an increase in trust; a sharp increase in the ability to verbalize feelings and to substitute such verbalization for aggressive acts; great increments in the ability to relate pleasurably with other members of the class; an increase in the development of personal competence; an increased appetite for books, trips, and other intellectually stimulating experiences; manifestations of personalized forms of expression in place of impersonal, constricted modes of responding. While it is probable that the children of this school came from more troubled and maladjusted families than those typically encountered in Head Start, it is instructive to observe how central was the role of non-cognitive elements in these teachers' analyses of the issues which impaired the functioning of their children. Cognitive growth was often depicted as advancing as a matter of course in response to the educational program once the obstacles to emotional involvement in school life were overcome.

In sum, the results of this study indicate that relatively few differences were found between these two groups of Head Start children and that the differences that were found did not form a consistent pattern. Larger differences were found between the middle-class group and the two Head Start groups. Some of these differences became larger at the conclusion of a year of preschool experience for all three groups. The latter findings are at variance with the expectations of those who have suggested that preschool experience for middle-class children serves mainly to foster certain desired social skills but leaves their normal and vigorous intellectual development unaffected. By the same token, this viewpoint has held that a year of preschool can serve to accelerate the retarded rates of development among disadvantaged children, thereby compensating for some of the deprivation they have previously experienced and reducing the gap between advantaged and disadvantaged.

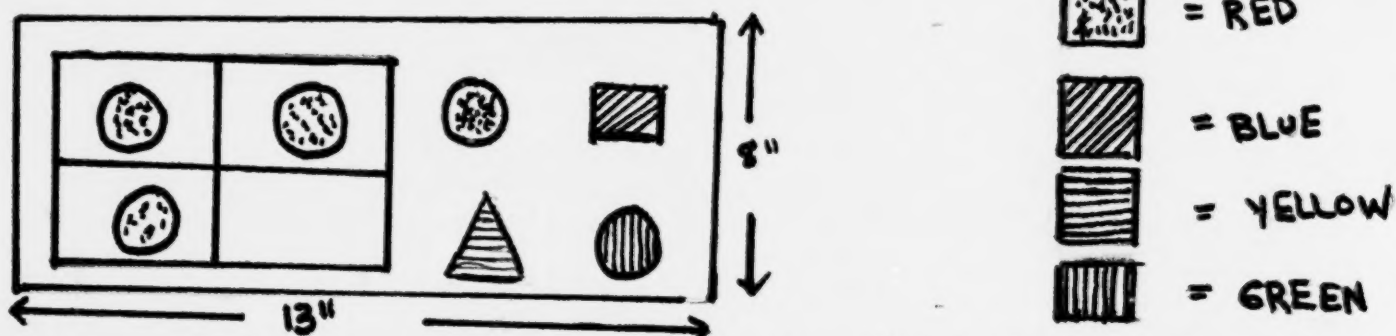
The present study, while far from definitive, and bedeviled by a number of methodological problems which have been noted, presents evidence that the privileged children seemed more able to benefit in demonstrable ways from a year of preschool education. It may be argued, contrary to the compensatory education hypothesis, that middle-class children, by virtue of the more fully developed cognitive base they bring to nursery school, are in a position to learn more from such an experience. While the data from this study are congruent with this latter hypothesis, it is important to note that the Hillcrest school was judged to offer a better educational program than that given by the two Head Start programs in the study. This comparative aspect of the study was marred by the fact that the advantaged and disadvantaged children were exposed to different preschool programs. In light of the fact that the Head Start programs included in the study were among the finest encountered among the initial survey of prospective field locations for the study, it would appear that the educational programs of most Head Start classes have not begun to approach the level of quality to be found in more established, privately operated nursery schools. More conclusive studies of the responsiveness of disadvantaged children to preschool experience will follow when the educational level of Head Start classes is raised.

APPENDIX

THE MATRIX TEST

The Matrix Test was devised to serve as a procedure for assessing classification, sorting and related cognitive skills associated with inferential reasoning. Based on a format used by Inhelder and Piaget (1964)* to study classification behavior in young children, it consists mainly of newly constructed items combined with a few devised by Inhelder and piaget. The test also resembles Raven's Progressive Matrices Test, but its format and content are more suited for use with young children -- it includes representational as well as abstract items, it requires a less abstract attitude, and it presents items individually, on separate cards (8" x 13") rather than in a booklet.

Each item of the test presenta a matrix of 2 x 2 or 2 x 3 squares in which all but one of the squares contain two-dimensional geometric figures or pictorial representations of familiar objects arranged so that the figures form some relationship to each other on the basis of their appearance, content or spatial position in the matrix. The subject is asked to find the figure missing from the empty square from among four alternatives presented alongside the matrix (see example below).



The subject merely has to point to the alternative that he believes to belong to the empty square. This format is simple to administer and has the additional advantage of easily communicating the essential task requirements. It avoids the kind of intricacy which the child may not understand, such as the request, in conventional sorting tasks, to "choose the objects which are alike." The conspicuousness of the empty square in the Matrix Test invariably communicates, even to the youngest child, that the appropriate figure must be found to fill it.

Content of the Matrix Test

The test is made up of 25 items selected from an original 44 items which were given to children in kindergarten through second grade. The 25 items have been chosen on the basis of their appropriateness to the preschool child. Four different classes of items may be distinguished: Perceptual Matching items,

*The Early Growth of Logic in the Child by Barbel Inhelder and Jean Piaget, New York: Harper and Row, 1964.

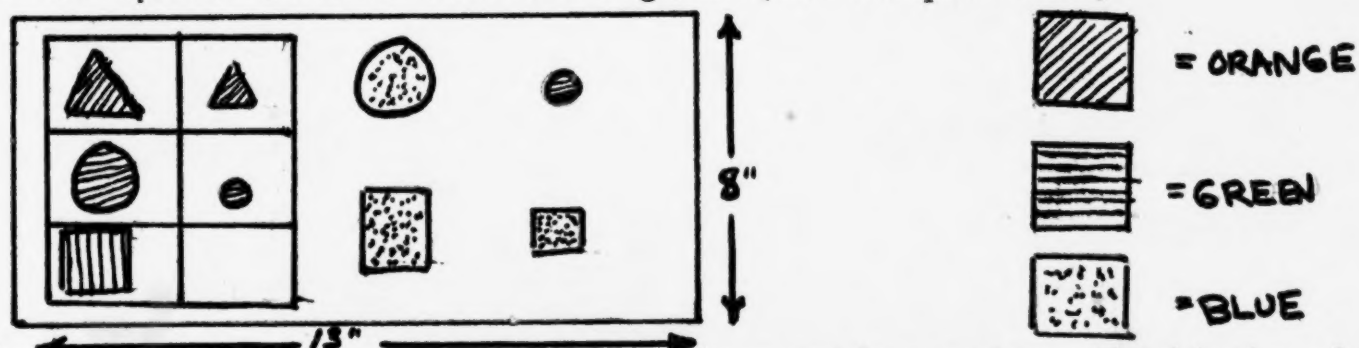
Class Membership items, One-Way Classification items, and Two-Way Classification items.

The Perceptual Matching items (N=4) present the easiest tasks. All four items consist of a 2 x 2 matrix in which the figures in all three occupied squares are identical. The task simply requires the child to find a fourth identical figure from among the alternatives. Both abstract and representational figures are included in this group. It should be emphasized that since these items merely require the child to find the matching figures, no abstraction or complex inference is entailed.

The Class Membership items (N=14) present a 2 x 2 matrix in which the three occupied squares contain different figures that have a common feature. In some of the items containing abstract figures, their color or form is the common feature; in others; the internal relationship of a combination of variables, such as size and color, constitutes the common element. Among the items presenting representational figures, the figures depict objects -- e.g., apple, pear, cherry -- that may be subsumed under some common category of classification -- e.g., fruit. These items vary in the degree of abstractness of the unifying category.

The One-Way Classification items (N=5) present 2 x 3 matrices (as well as some 2 x 2 matrices) of abstract or representational figures in which all the members of the vertical arrays (columns) or horizontal arrays (rows) are the same. Thus the identity of the missing figure is given by its column or row membership.

The Two-Way Classification items (N=2) present 2 x 3 matrices in which the row and column membership, in combination, determine the nature of the missing figure. Thus, whereas all the members of the same row or column of the One-Way Classification items are identical, in the Two-Way Classification items, no two squares contain identical figures (see example below).



The Matrix Test is being administered in a uniform order. The blocks of the items are presented in toto, in the same order in which the four groupings have been presented here. During administration of the test, the child is not told of the transitions in task requirement of the blocks of items presented to him and there is no attempt to probe the child's response to an item. Variations in the sequence and mode of presentation of the items, and experimentation with a form of inquiry, are currently being contemplated.

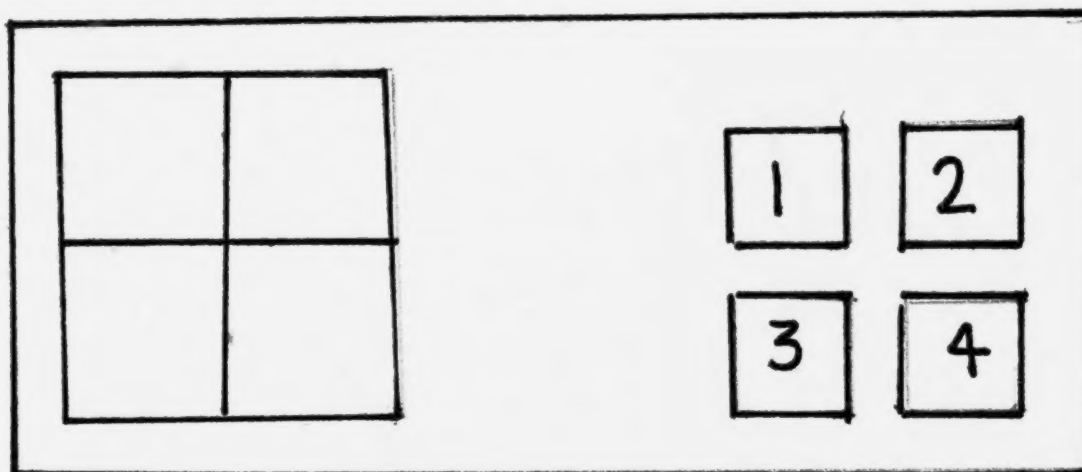
Instructions

Begin by explaining to the child that you have some pictures to show him and

a game to play with him. Take the first card and after covering up the alternatives place it in front of the child and say, "Look at these boxes. You can see that there is something in this box (pointing), something in this box (pointing), something in this box (pointing), but nothing in this box (pointing)." Uncover the alternatives and say, "Let's find the one that belongs here (pointing to empty box). Point to the one which belongs here with the others." On subsequent cards one need only remind the child to "Point to the one that belongs with the others." When each card is presented, the alternatives should be covered for four or five seconds.

Recording Responses

Each alternative is assigned a number according to its position on the card (see below). All responses should be recorded by number, i.e., the tester should place a check in the appropriate column (see record sheet).



KALEIDOSCOPE SITUATION

This technique is included as a measure of the child's curiosity and exploration.* A kaleidoscope is the only object on the table when the child enters the room. The examiner, who in the eyes of the child is attending to some paper work, observes and records the child's reactions for a period of two minutes. He maintains a neutral stance (e.g., if the child asks "Can it open?" the examiner might say, "You may play with it in the way you like.") If, after 30 seconds, the child has made no attempt to touch the kaleidoscope the examiner says "Would you like to play with that (indicating the kaleidoscope) while I finish with what I'm doing?" If, after a further 30 seconds the child still has not touched the kaleidoscope, the examiner picks up the kaleidoscope and offers it to the child, repeating his previous cue. If the child still doesn't respond, the examiner waits 60 seconds before terminating the task.

Analysis of the record is on a point system and yields an object curiosity score (see below).

Point System for Object Curiosity Score

Maximum possible score: 18

One point credit for each numbered item below:

- | | |
|----------------------------|--|
| Initial Reaction: | 1. Reacts immediately on own |
| Span of Involvement: | 2. Spends two full minutes exploring without I's prompting (cannot be credited if #1 not credited) |
| Questions: | 3. General questions of identification ("What is this?" etc.)
4. Other questions about object ("Can it open?" etc.) |
| Object Manipulation: | 5. Turns object different ways
6. Turns moving part
7. Shakes object
8. Looks in object
9. Other |
| Spontaneous Verbalization: | 10. Comments on what is seen in kaleidoscope |

Two points for each numbered item below:

- | | |
|---------------------------|--|
| Novelty-Seeking Behavior: | 11. Turns moving part in order to produce visual change (intent must be clear) |
|---------------------------|--|

*See Patricia Minuchin, "Processes of Curiosity and Exploration in Preschool Disadvantaged Children." Final report of study, OEO Contract No. OEO-2403.

12. Shakes object to produce visual change
13. Changes object angle to produce change (up, down, to light, etc.)

Extension of Object Use:

14. Incorporates into play; uses object in different ways (as flashlight, camera, etc.)

The observer uses the attached checklist wherever possible and, in addition, makes written observations when appropriate.

KALEIDOSCOPE
(Checklist)

Child's name _____
Center _____
Class (a.m. or p.m.) _____

Sex _____
Tester _____
Date _____

Check if
Applicable

- | | | |
|--|-------|--|
| 1. Reacts immediately on own | _____ | |
| 2. Spends two full minutes exploring without I's prompting (cannot be credited if #1 not credited) | _____ | |
| 3. General questions of identification ("What is this?" etc.) | _____ | |
| 4. Other questions about object ("Can it open?" etc.) | _____ | |
| 5. Turns object different ways | _____ | |
| 6. Turns moving part | _____ | |
| 7. Shakes object | _____ | |
| 8. Looks in object | _____ | |
| 9. Other | _____ | |
| 10. Comments on what is seen in kaleidoscope | _____ | |

Items 1-10 (one
point each) Total =

- | | | |
|--|-------|--|
| 11. Turns moving part in order to produce visual change (intent must be clear) | _____ | |
| 12. Shakes object to produce visual change | _____ | |
| 13. Changes object angle to produce change (up, down, to light, etc.) | _____ | |
| 14. Incorporates into play; uses object in different ways (as flashlight camera, etc.) | _____ | |

Items 11-14 (two
points each) Total =

(Check one):

Child responded without cues	_____
Child responded after first cue	_____
Child responded after second cue	_____
Child did not respond	_____

Total Score

Individual Child Observations

There were mid-year observations of individual children in which the observer recorded the child's activities during a given period. Each child was observed in two cycles. Each cycle consists of two 15-minute periods. The second observation cycle was on a different day and covered different activities than the first. The following dimensions guided the observations:

- A. Purposefulness of child's behavior -- the degree to which behavior is purposeful or aimless or seemingly random.
- B. Mode of mastery re materials -- child's ability to use the materials and tools at his disposal, the degree to which he is process vs. end product oriented.
- C. Mastery of school environment -- child's awareness of where objects belong and can be found, the degree to which he knows his way around in the classroom, playground, etc.
- D. Frustration tolerance -- the sources of frustration for the child (e.g., enforced stillness, inability to communicate needs) and his response (e.g., hitting, crying, withdrawal).
- E. Impulse control -- child's ability to control aggressive reactions, to inhibit crying and other emotional outbursts, physical hurt, destruction of property, etc.
- F. Mode of social relationships with other children -- child's characteristic approach to other children (verbal and physical) and his characteristic response to the approaches of others (verbal and physical).
- G. Curiosity -- the extent to which the child expresses and sustains interest in novel events, inquires about how things work or why certain things occur, his readiness to listen to explanations and to search for understanding.
- H. Strength of self -- his competence, awareness of what he wants, and ability to assert himself.

Functional Category System for the Analysis
of Spontaneous Interpersonal Language of
Preschool Children

This is a measure which is being developed to categorize the spontaneous interpersonal communications of preschool children in their classrooms. It will be used to assess differences in language as a communicative (interpersonal) process among preschool children, as well as their ability to use language as an intellectual (intrapersonal) process.

Verbatim samples of interpersonal communication are recorded for each child during free play in 12 three minute periods. The sample will consist of equal numbers of boys and girls who are matched for high, medium and low IQ (109+, 91-108, and 90- respectively). Each spontaneous, interpersonal verbal statement semantically defined, is scored on a category system. There are nine major functional categories in this system:

I. Expressive. Statements and exclamations which serve an expressive function, both negative ("I hate this") and positive ("Yummy!").

II. Desire Implementing. Statements to another which implement a desire of S. This category includes specific requests, e.g. for milk, as well as requests for reassurance, direction and guidance, permission to do something, and requests for attention, by clowing, etc.

III. Rights Implementing. Statements which implement a right involving an object, territory, a turn, or a role ("I want that because it's mine," "Get out of my corner").

IV. Egocentric-Pride Enhancing. Statements in this category are usually of two kinds: (a) those having to do with power or competence ("Look what I did") and potency or prestige ("Look how tall I am"), and (b) those having to do with a good-bad evaluation ("I'm right," "I know where the blocks should go").

V. Me-Too. Statements in which the S joins himself to others by self-referral or self-inclusion ("I saw that too," "I want to play too").

VI. Joining. Statements in which the S joins with another by a simple uniting statement ("Hello," "I like you") or by engaging the other in joint action ("Let's play together").

VII. Collaborative. Statements which create or maintain a role differentiated collaboration, where S participates with one or more children in a project, discussion or game, with each child playing a part in a joint endeavor ("Now you be the baby" in dramatic play).

VIII. Reporting. Neutral statements about the world, the self, etc., which are not need-fulfilling or ego-enhancing ("It's raining" "Look at the guinea pig").

IX. Learning Implementing. Statements in this category function to implement the desire to learn. They fall into two main sub-categories: (a) searching for information ("What does that say?"), and (b) reitcrating and reinforcing previous learning ("Babies don't have teeth, right?").

Each of these nine categories is divided into numerous sub-categories (over 100 in all) which further differentiate the functions of statements.